

## A FRAMEWORK FOR ONTOLOGY DEVELOPMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY EXPERTS USING THESAURUS, ASSOCIATION FOR COMPUTING MACHINERY TAXONOMY AND DOMAIN EXPERTS APPROACHES

AKMAL ARIS  
JUNAIDAH MOHAMED KASSIM  
HANA YASMEIN ISHAK  
JUHANA SALIM  
SHHRUL AZMAN MOHD NOAH

### ABSTRACT

*A search for experts is a search for individuals with skills and knowledge in a particular field. Early studies studied profiling of experts in the expert search application by using keyword query. The use of concept-based query that we highlight is able to match experts and expertise to the searched field. Existing expert search systems perform search merely based on the experts' publication information. There are no relations to other information that can act as support resources to relate them to related fields. The technique to search for expert based on publication information narrows the search scope. There are other aspects that are related to the experts such as education, teachings, supervisions, consultations, conferences, community services and professional affiliations. The use of ontology that is based on concept can also help to expand the search results as well as to link to other relations. The purpose of this research is to design a framework for ontology development of Information, Communication and Technology (ICT) experts, to be used in the development of ICT expert directory. This framework is developed based on the critical evaluation of the literatures focusing on the following areas: experts search systems, experts profiling and ontology in the experts search system. The development of the framework involves extraction processes from web resources, for example, resume, individual websites, organization websites and research centers websites, to build a glossary of word index that characterizes experts. This glossary index becomes the basis for ontology development. In this research, the ontology development integrates the use of thesaurus, Association of Computing Machinery (ACM) taxonomy and ICT domain experts. The ontology that had enriched the glossary index provides the semantics ability in the ICT expert directory. There are three components in the framework; they are extraction of data, development of ontology, and validation of ICT expert ontology by domain experts. The development of this framework can guide the development of concept-based ontology that is relevant and accurate to user's needs.*

## KERANGKA PEMBANGUNAN ONTOLOGI PAKAR TEKNOLOGI MAKLUMAT DAN KOMUNIKASI MENGGUNA PENDEKATAN TESAUROS, TAKSONOMI ASSOCIATION FOR COMPUTING MACHINERY DAN PAKAR DOMAIN

### ABSTRAK

Pencarian pakar adalah pencarian individu yang mempunyai kemahiran dan pengetahuan dalam sesuatu bidang. Kajian lampau mengkaji penggunaan profil pakar dalam aplikasi pencarian pakar menggunakan kata kunci sebagai kueri. Penggunaan kueri berasaskan konsep yang diketengah mampu mendapat padanan pakar atau kepakaran yang berkaitan dengan bidang yang dicari. Sistem pencarian pakar kini hanya melakukan carian berdasarkan maklumat penerbitan pakar sahaja lantas menyebabkan tidak ada hubungan dengan maklumat lain yang boleh menjadi sumber sokongan bagi menghubungkan kepada bidang yang berkaitan. Kaedah pencarian pakar berdasarkan maklumat penerbitan mengencil skop carian pakar kerana terdapat aspek lain yang berkaitan dengan pakar seperti pendidikan, pengajaran, penyeliaan, perundingan, persidangan, khidmat masyarakat, kelayakan dan badan

profesional. Penggunaan ontologi yang berasaskan kepada konsep dapat membantu mengembang hasil carian serta menunjukkan hubungan yang berkaitan. Kajian ini bertujuan mereka bentuk satu kerangka bagi pembangunan ontologi pakar teknologi maklumat dan komunikasi (TMK) yang diguna dalam pembangunan direktori pakar TMK. Kerangka dibina berdasarkan kepada penilaian kritis susastera dalam bidang sistem pencarian pakar, profil pakar, dan ontologi dalam sistem carian pakar. Kerangka pembangunan ontologi pakar TMK melibatkan pengestrakan maklumat pakar dari sumber web seperti resume, laman web individu, laman web organisasi dan laman web pusat penyelidikan bagi membina indeks glosari perkataan yang mencirikan pakar. Indeks glosari yang dibina hasil daripada pengestrakan dijadikan asas bagi membangun ontologi. Kajian ini mengintegrasikan penggunaan tesaurus dan taksonomi Association of Computing Machinery (ACM) dalam pembangunan ontologi pakar yang melibatkan pakar domain dalam bidang TMK. Ontologi dapat memperkaya indeks glosari yang diguna dalam sistem direktori pakar yang menyediakan keupayaan semantik. Kerangka pembangunan ontologi pakar TMK yang dicadangkan merangkumi tiga bahagian iaitu pengekstrakan data, pembangunan ontologi dan pengesahan ontologi pakar TMK oleh pakar domain. Kerangka pembangunan ini dapat memandu pembangunan ontologi berasaskan konsep yang relevan dan tepat mengikut kehendak pengguna.

Kata kunci: Pakar, profil pakar, sistem pencarian pakar, ontologi dan kerangka kerja pembangunan ontologi.

## PENGENALAN

Pakar merupakan individu yang berpengetahuan dan berkemahiran dalam sesuatu bidang. Seseorang individu atau sesebuah organisasi yang ingin mendapat khidmat nasihat bagi menangani sesuatu isu atau masalah perlu mengetahui dan mencari pakar yang boleh dirujuk. Justeru, persoalan “Siapakah pakar dalam bidang X?” perlu dijawab. Maklumat pakar boleh didapati dalam web persendirian, web organisasi, rangkaian sosial dan medium perantaraan yang lain. Sekiranya teknologi web semantik diaplikasi ke dalam teknologi web yang terkini, penyampaian maklumat atau penghubungan pakar dengan masyarakat menjadi pantas dan keputusan carian maklumat pakar menjadi relevan.

Bagi menyokong teknologi web semantik dalam pencarian pakar, ontologi diguna sebagai komponen utama dalam sistem pencarian pakar. Penggunaan ontologi dalam carian pakar dibukti oleh Punnarut dan Sriharee (2010) mampu menyelesaikan permasalahan yang wujud dalam pencarian pakar. Ontologi menyebabkan hasil pencarian tidak lagi tertumpu pada senarai pakar sahaja sebaliknya segala maklumat yang berhubung dengan pakar boleh diperolehi dengan mudah. Selain daripada itu, ontologi juga dapat menyelesaikan keputusan carian yang tidak relevan kerana istilah dan konsep yang diguna oleh pengguna sering tidak bertepatan dengan keperluan maklumat mereka. Justeru, sistem pencarian pakar yang diintegrasikan dengan ontologi merupakan inisiatif bagi mendapat padanan pakar yang relevan.

Pembangunan ontologi untuk profil pakar dalam sistem pencarian pakar adalah aplikasi yang khusus bagi domain tertentu. Analisis terhadap kajian lampau mendapati pembangunan ontologi bagi profil pakar teknologi maklumat dan komunikasi (TMK) kurang diberi perhatian lantas memotivasi kajian ini dijalankan yang bertujuan menghasilkan sebuah kerangka pembangunan ontologi pakar TMK.

Kajian tentang pencarian pakar perlu mengetahui maklumat yang diperlu bagi memprofil maklumat pakar. Contohnya, maklumat pakar boleh diekstrak dari sumber web seperti resume, laman web individu dan laman web organisasi. Indeks glosari yang dibina hasil daripada pengestrakan dijadikan asas bagi membangun ontologi. Penyelidikan terdahulu menumpu kepada pencarian seseorang pakar tanpa melihat dari aspek memperkaya maklumat pakar tersebut.

Pembangunan ontologi yang melibatkan pakar domain (Xing et al., 2009 dan Siti Farhana et al., 2014) dapat memperkaya maklumat seseorang pakar dengan meperluas pencarian pakar. Penglibatan pakar domain dapat memberi profil topikal tentang pakar dan profil sosial serta mengenal pasti perkara lain dan pihak lain yang ada kaitan dengan pakar. Namun, kebergantungan kepada pakar domain sahaja boleh mengekang pembangunan ontologi dan pengetahuan pakar domain tidak merangkumi keseluruhan domain yang dikaji.

Tesaurus dan taksonomi diguna oleh penyelidik terdahulu bagi memperkaya perbendaharaan kata dan memberi keputusan carian yang relevan. Pendekatan baharu dalam metod pembangunan ontologi pakar adalah mengguna tesaurus dan taksonomi seperti taksonomi Association for Computing Machinery (ACM) berbanding dengan hanya bergantung kepada pakar domain.

Objektif kajian ini ialah mereka bentuk kerangka pembangunan ontologi domain pakar TMK mengguna pendekatan taksonomi, tesaurus dan pakar domain. Kerangka ini mengandungi tiga komponen iaitu pengekstrakan sumber pakar dan kepakaran, pembangunan ontologi dan pengesahan kandungan ontologi.

## KEPERLUAN PEMBANGUNAN ONTOLOGI PAKAR

Individu yang mempunyai kepakaran bagi menyelesaikan masalah merupakan satu keperluan genting bagi sesebuah organisasi. Namun proses mencari kepakaran sering berhadapan dengan masalah lantaran padanan dokumen dan pakar yang bersesuaian yang tidak tepat (McDonald, 2001).

### ONTOLOGI PAKAR

Ontologi merupakan satu aspek yang menjadi tumpuan pengkaji lepas berhubung dengan sistem carian pakar. Pendekatan mengguna ontologi bertujuan mengenal pasti pakar yang sesuai dengan kata kunci yang dicari berserta maklumat tambahan yang berkaitan dengan hasil carian yang diberi oleh sistem pencarian pakar.

Liu et al. (2007) mencadangkan penggunaan ontologi pakar bagi mengintegrasikan pelbagai penunjuk kepakaran dari sumber data yang pelbagai. Penggunaan ontologi domain dalam pencarian profil kepakaran berdasarkan kata kunci konsep adalah bagi menggantikan pencarian profil berdasarkan kata kunci. Kajian Liu et al. (2007) bertujuan meningkatkan peluang mencari pakar yang relevan dan menyokong pengguna memilih pakar yang sesuai dengan menyediakan maklumat yang terperinci setiap pakar. Menurut Liu et al. (2007), ontologi domain dibina bagi menyimpan kata kunci konsep dan hubungan antara domain yang berkaitan dengan koleksi dokumen. Pendekatan adalah bagi mengatasi kelemahan dalam pencarian pakar tradisional yang lazim memerlukan pangkalan data kemahiran atau kepakaran yang khusus. Tinjauan turut dijalankan terhadap metod pembangunan ontologi sedia ada bagi memberi idea dalam merangka pendekatan baharu pembangunan ontologi pakar.

Tinjauan mendapati semua metod pembangunan ontologi bergantung sepenuhnya kepada pakar sesuatu domain. Xing et al. (2009) menyatakan kepentingan pakar domain adalah perlu dalam pembangunan ontologi bagi menerangkan sesuatu bidang yang khusus dengan jelas tentang konsep yang tidak terdapat dalam tesaurus.

### PAKAR DOMAIN

Tinjauan ke atas kajian tentang pembangunan ontologi domain mendapati penglibatan jurutera pengetahuan adalah perlu bagi membangun sesuatu ontologi domain kerana dapat berfungsi sebagai sumber yang boleh dipercayai (Ensan & Du, 2007). Keperluan sedemikian disokong oleh Carvalho et al. (2004), Simperl dan Mochol (2006) dan Omoronyia et al. (2010) yang menyatakan pakar domain adalah individu yang dapat menyumbang pengetahuannya dalam sesuatu domain yang ingin dimodelkan dalam bentuk ontologi. Sumbangan ini adalah dalam bentuk istilah abstrak berdasarkan pengalaman dan pengetahuan (Xing et al., 2009; Fonou-Dombeu & Huisman, 2011, Siti Farhana et al., 2014). Pembangunan ontologi kandungan kursus sememangnya wajar melibatkan pakar domain dan jurutera ontologi sebagaimana yang dilakukan oleh Boyce dan Pahl (2007).

Pakar domain dalam kajian ini melibatkan proses pengesahan istilah yang diekstrak dari buku teks untuk subjek kursus sebagai sumber kepada ontologi yang dibangun. Tugas pakar domain pula tidak terbatas kepada peringkat awal sahaja, sebaliknya penglibatan semasa proses penilaian turut dapat membantu menghasilkan konsep yang tepat dalam ontologi.

Bagi memasti istilah yang diguna memenuhi piawaian, Xing et al. (2009) menggunakan tesaurus meskipun tesaurus tidak mampu menjelaskan sesuatu bidang dengan khusus. Menurut pakar domain, penambahan istilah baharu perlu dilakukan bagi menjelaskan sesetengah konsep yang penting tetapi tidak wujud dalam tesaurus.

## TAKSONOMI

Taksonomi ialah istilah kajian tentang prinsip, peraturan dan amalan dalam pengkelasan. Daconta et al. (2003) mendefinisi taksonomi dari perspektif teknologi adalah pengkelasan maklumat dalam bentuk hierarki mengikut perwakilan hubungan dalam dunia sebenar. Pengkelasan taksonomi menggunakan hubungan *father-son* (*generalization, is-a* atau *type of*).

Pengkelasan merupakan salah satu bentuk taksonomi (Dacota, 2003). Menurut Juhana et al. (2007) terdapat tiga jenis skema pengkelasan umum iaitu Pengkelasan Perpuluhan Dewey, Pengkelasan Perpuluhan Universal (UDC) dan Pengkelasan Library of Congress (LCC) yang diguna oleh perpustakaan bagi mengkatalog koleksi perpustakaan. Bagi mengecil skop kepada sesuatu bidang kerja, terdapat skema subjek khusus yang direka untuk kumpulan pengguna khas. Terdapat juga skema pengkelasan yang tertumpu pada satu topik khas seperti skema pengkelasan industri dan perniagaan. Skema pengkelasan Standard Industrial Classification (SIC), skema North American Industry Classification (NAIC) dan skema pengkelasan United Nations Standard Products dan Service Code (UNSPRC) merupakan skema utama skema pengkelasan industri dan perniagaan. Penggunaan skema subjek khusus berpotensi menyediakan struktur dan istilah yang lebih hampir dengan disiplin dan terkini berbanding skema umum (Juhana et al., 2007).

Penggunaan taksonomi ACM atau sistem pengkelasan ACM adalah bagi mewakili kepakaran dalam bidang Sains Komputer (Mirkin et al., 2007). Taksonomi ACM diguna bagi mengorganisasi topik penyelidikan mengikut organisasi. Selain daripada itu, taksonomi ACM boleh diguna bagi tujuan berikut:

1. mendapat gambaran keseluruhan tentang pembinaan subjek terutama bidang sains komputer dalam sesebuah organisasi
2. menyusun atur organisasi berdasarkan taksonomi ACM
3. merancang penstrukturan penyelidikan dan pelaburan dalam bidang sains komputer.

Menurut Schult dan Spiliopoulou (2006), taksonomi ACM diguna dalam cadangan kata kunci, pengkategorian dan carian dalam pengkelasan dokumen. Namun terdapat beberapa kelemahan pada taksonomi ACM iaitu kandungan taksonomi ACM tidak mencakupi topik terkini. Oleh yang demikian, sumber rujukan alternatif bagi mengatasi masalah ini diperlu.

## TESAURUS

Jumlah maklumat web meningkat dari sehari ke sehari. Enjin carian sedia ada menghadapi masalah seperti ketidaksesuaian penggunaan kata kunci ketika melakukan carian dan kata kunci yang dimasukkan adalah pendek (Chen et al., 2003). Purata panjang kueri yang lazim diguna semasa melakukan pencarian ialah dua patah perkataan (Wen et al., 2001). Oleh kerana kueri yang dimasukkan pendek, maka keputusan hasil carian adalah tidak memuaskan. Bagi meningkatkan hasil carian dengan dokumen yang relevan, kueri carian perlu dikembangkan.

Menurut Stefanov (2003), penggunaan tesaurus dalam bidang capaian maklumat melibatkan proses pembinaan automatik, reka bentuk antara muka pengguna, mekanisme

capaian, dan seni bina capaian. Terdapat tiga isu asas yang berkaitan dengan tesaurus dalam capaian maklumat (Jing & Croft, 1994) iaitu pembinaan tesaurus, capaian tesaurus dan penilaian kandungan tesaurus. Thesaurus Roget's dan Wordnet adalah umum dan berdasarkan perkataan yang mengandungi hubungan seperti antonim dan sinonim. Ini berbeza dengan INSPEC, Library Congress Subject Heading (LCSH) dan Medical Subject Heading (MeSH) yang merupakan tesaurus yang berorientasi capaian maklumat dan berasaskan frasa.

Selain daripada tesaurus umum, terdapat tesaurus khusus seperti tesaurus The Institute of Electrical and Electronics Engineers (IEEE) yang menjurus kepada kejuruteraan, teknikal, dan istilah sains serta terma khusus dalam komuniti IEEE. Tesaurus IEEE adalah perbendaharaan kata yang terkawal terdiri daripada 9000 deskriptif kejuruteraan, teknikal, dan istilah sains serta terma khusus komuniti IEEE (<http://www.ieee.org>). Tesaurus tersebut mewakili penggunaan istilah dan kosa kata yang diambil daripada transaksi IEEE dan artikel jurnal, kertas persidangan, dan piawai. Skop dan struktur tesaurus IEEE menyatakan bidang kejuruteraan dan saintifik yang terdiri daripada persatuan, majlis, dan komuniti IEEE sebagai tambahan kepada perkhidmatan teknologi IEEE. Rajah 1 menunjukkan contoh tesaurus IEEE.

<b>Knowledge management</b>	
<b>UF:</b>	Intellectual capital
<b>BT:</b>	Computer applications Management
<b>RT:</b>	Competitive intelligence Information management Knowledge engineering Management information systems Semantic Web
<b>NT:</b>	Knowledge transfer

RAJAH 1. Contoh tesaurus IEEE

Sumber: <http://www.ieee.org/documents/pdfieethes05nov09.pdf>

Penggunaan tesaurus IEEE memudah penyelidik memperkaya istilah dalam bidang TMK. Sebagai contoh, untuk tajuk *Knowledge management*, didapati *Use For* (UF) adalah *intellectual capital*, *Broad Term* (BT) adalah *computer application management*, *Related To* (RT) adalah *competitive intelligence*, *information management*, *knowledge engineering*, *management information system* dan *semantic web* dan *Narrow Term* (NT) adalah *knowledge transfer*. Mengguna tesaurus IEEE, istilah *knowledge management* boleh dikembang kepada beberapa istilah lain. Kajian penggunaan tesaurus web IEEE dilakukan oleh Sattar dan Saeed (2002) melibatkan penyelidik menggabung sistem pengelasan DDC dan tesaurus web IEEE bagi membina taksonomi bagi domain sains komputer. Tesaurus ini diguna sebagai sumber maklumat utama bagi pengelasan bidang elektronik dan komputer.

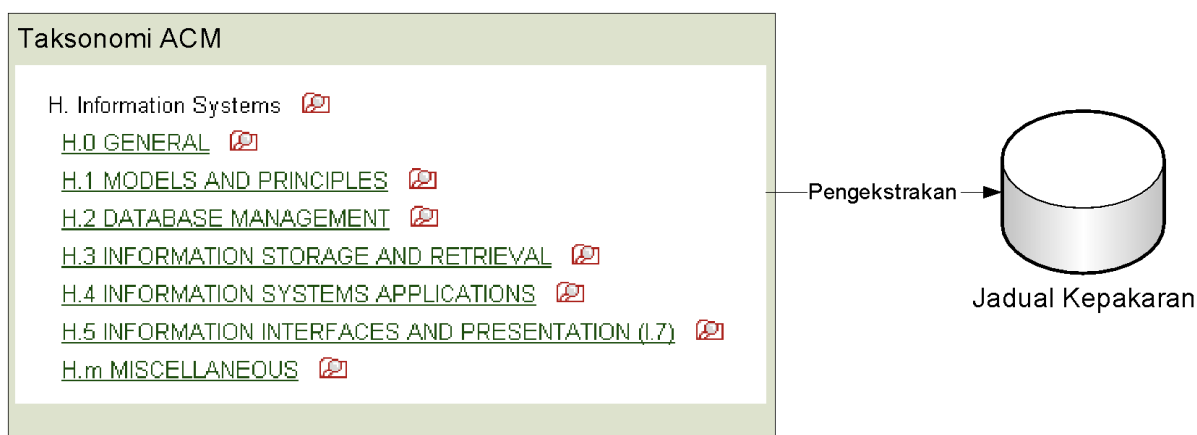
#### PEMBANGUNAN ONTOLOGI PAKAR

Pembangunan ontologi pakar TMK melibatkan beberapa langkah dalam membangun ontologi. Langkah yang dikenal pasti ialah pengenalanpastian sumber data dalam membangun ontologi, pengekstrakan sistem pengelasan pengkomputeran ACM, pengembangan istilah dari sumber eksplisit yang berautoriti dan pembangunan ontologi ke dalam alat pembangunan ontologi seperti *Protege*.

Sebelum ontologi dibangun, sumber data yang diperlu dikenal pasti. Sumber data yang dipilih terdiri daripada kepakaran bidang TMK dan sumber pakar. Bagi sumber kepakaran, sistem pengkelasan pengkomputeran ACM menjadi sumber utama bidang kepakaran. Sumber ini dipilih berdasarkan kepada kajian Punnarut dan Sriharee (2010) yang mengguna sumber yang sama bagi membangun ontologi klasifikasi kecekapan. Sumber data kedua pula ialah maklumat subjek kajian iaitu pakar. Liu et al. (2007) mendapati sumber maklumat penerbitan diguna sebagai sumber maklumat pakar. Bagaimanapun terdapat kekurangan dari segi hubungan dengan pakar lain apabila hanya satu sumber sahaja diguna. Oleh yang demikian, kajian ini mengguna sumber tambahan iaitu resume pakar sebagai sumber maklumat.

PENGEKSTRAKAN SISTEM PENKELASAN PENKOMPOTERAN ACM DAN MAKLUMAT PAKAR

Dalam fasa ini, terdapat dua proses utama iaitu pengekstrakan sistem pengkelasan pengkomputeran ACM sebagai maklumat kepakaran dan pengekstrakan maklumat pakar dari resume. Pertama, sistem pengkelasan pengkomputeran ACM diekstrak bagi mendapat maklumat berkenaan dengan kepakaran dalam bidang TMK. Sumber sistem pengkelasan pengkomputeran ACM dimuat turun dari web portal ACM (<http://www.acm.org/about/class/1998/>) dalam bentuk format Extensible Markup Language (XML). Semasa proses pengekstrakan dijalankan (rujuk Rajah 2), segala maklumat yang terkandung dalam fail tersebut disimpan dalam jadual kepakaran.



RAJAH 2. Proses pengekstrakan sistem pengkelasan pengkomputeran ACM

Semasa proses pengekstrakan dijalankan, sistem mengasing setiap kategori berdasarkan huruf yang mewakili setiap kategori. Proses pengekstrakan kedua melibatkan maklumat pakar. Berdasarkan kajian Liu et al. (2007), pelbagai sumber maklumat boleh diekstrak bagi membangun ontologi pakar dan resume merupakan salah satu sumber tersebut. Dalam membentuk profil pakar, kajian ini mengguna beberapa maklumat seperti penerbitan, kepakaran, penyelidikan, perundingan, persidangan, dan reka cipta bagi memperoleh profil kepakaran pakar yang kemudiannya diekstrak ke dalam jadual pakar.

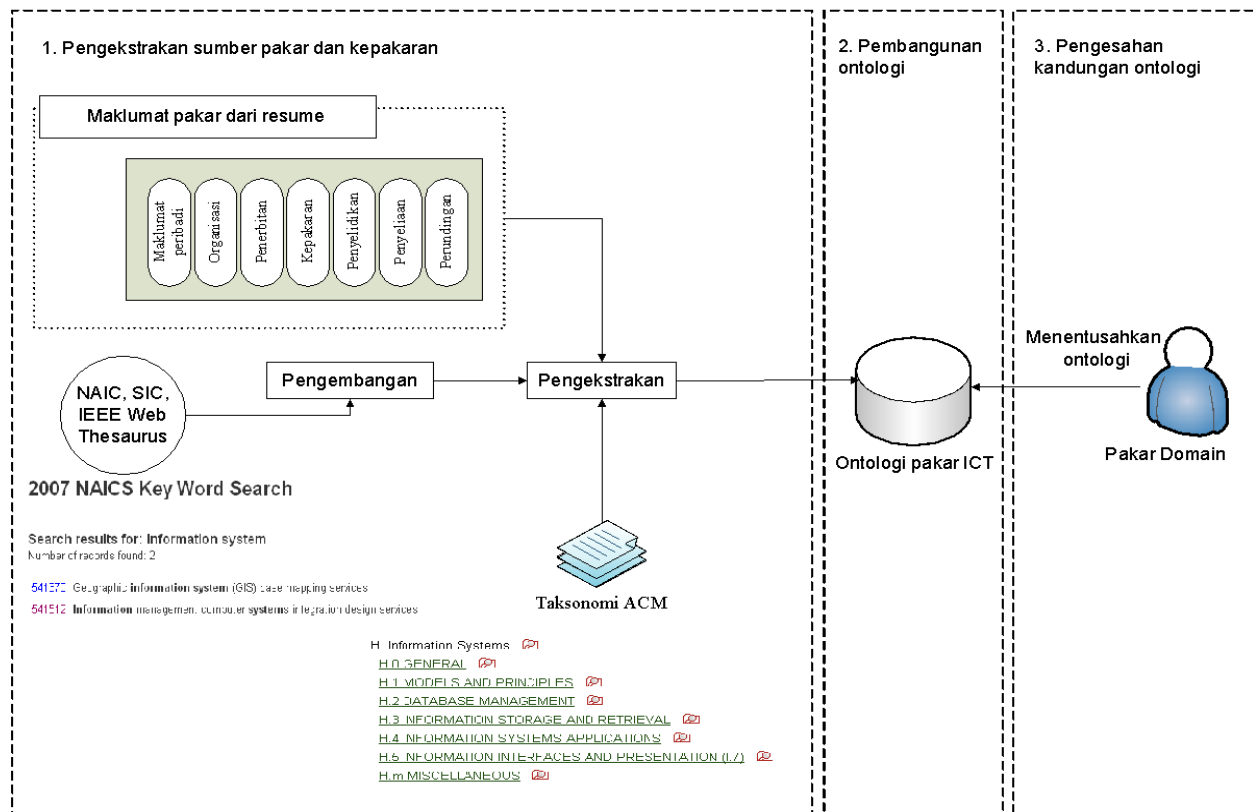
Setelah proses pengekstrakan kedua-dua data dijalankan, proses pengembangan istilah bagi bidang kepakaran pula dilakukan. Proses ini selari dengan kajian Velardi et al. (2001) yang mendapati ontologi boleh diperkaya dengan definisi dan hubungan bagi setiap istilah.

## HASIL DAN PERBINCANGAN

Kerangka pembangunan ontologi bagi pembangunan prototaip sistem direktori pakar TMK adalah seperti Rajah 3. Kerangka pembangunan ontologi pakar TMK yang dibangun boleh dibahagi kepada tiga komponen mengikut proses iaitu:

### PENGEKSTRAKAN SUMBER PAKAR DAN KEPAKARAN

Terdapat dua proses pengekstrakan yang melibatkan dua sumber data iaitu sumber kepakaran bidang TMK dan sumber maklumat pakar. Bagi sumber kepakaran, sistem pengkelasan pengkomputeran ACM diguna. Sebagaimana yang diguna oleh Punnarut dan Sriharee (2010) dalam pembangunan ontologi pengkelasan kecekapan. Sumber data kedua ialah maklumat tentang pakar. Meskipun Liu et al. (2007) hanya mengguna sumber maklumat penerbitan dalam kajiannya, kajian ini memilih sumber tambahan iaitu resume pakar. Tindakan sedemikian diambil kerana apabila hanya satu sumber sahaja diguna maka terdapat kekurangan dari segi hubungan dengan pakar lain. Oleh yang demikian resume kakitangan akademik Fakulti Teknologi dan Sains Maklumat (FTSM), Universiti Kebangsaan Malaysia (UKM) dipilih sebagai sumber tambahan bagi memperoleh maklumat tentang pakar.



RAJAH 3. Kerangka Pembangunan Ontologi Pakar TMK

Setelah proses pengekstrakan kedua-dua data, proses pengembangan istilah bagi bidang kepakaran pula dilakukan. Proses ini selari dengan kajian Velardi et al. (2001) yang mendapati ontologi boleh diperkaya dengan definisi dan hubungan bagi setiap istilah. Proses seterusnya ialah pengembangan istilah mengguna sumber eksplisit berautoriti. Data kepakaran yang diekstrak diambil dan dikembangkan mengguna sumber eksplisit seperti skema pengkelasan North American Industry Classification (NAIC), Standard Industrial Classification (SIC) dan tesaurus web IEEE. Menurut Juhana et al. (2009), dengan mengguna skema pengkelasan, capaian kepada sumber web menjadi cekap kerana sesebuah laman web yang mengguna skema

pengkelasan bersifat teratur. Kelebihan skema pengkelasan ini ialah laman web menjadi berstruktur, kandungan laman web dapat disemak dengan cepat, dapat meluas atau mengecil skop carian, memberi konteks kepada carian dan memudah pembahagian dan pengurusan pangkalan data. Kelebihan yang terdapat pada sistem pengkelasan mendorong kajian ini menggunakannya dalam proses pengembangan istilah kepakaran TMK yang diekstrak sebelum ini. Hasil proses pengestrakan yang dibuat merupakan profil pakar yang mempunyai hubungan di antara pakar dan domain kepakaran.

#### PEMBANGUNAN ONTOLOGI

Ontologi pakar TMK dibangun mengguna hasil proses pengestrakan yang dilakukan. Pembangunan ontologi pakar TMK melibatkan dua fasa, iaitu:

1. Fasa mengenal pasti domain kajian pembangunan ontologi pakar TMK  
Kajian mengenal pasti domain kajian iaitu memilih pakar bidang TMK seperti kajian Punnarut dan Sriharee (2010) dan Liu et al. (2007). Berdasarkan kajian Punnarut dan Sriharee (2010), sistem pengkelasan pengkomputeran ACM dijadikan sebagai sumber rujukan utama bagi bidang TMK dalam pembangunan ontologi kerana sistem pengkelasan pengkomputeran ACM mempunyai piawaian yang khusus dan memudah penyelidik mendapat maklumat bidang kepakaran.
2. Fasa pembangunan ontologi pakar TMK  
Beberapa langkah dalam membangun ontologi dikenal pasti seperti sumber data dalam membangun ontologi, pengestrakan sistem pengkelasan pengkomputeran ACM, pengembangan istilah dari sumber eksplisit yang berautoriti dan pembangunan ontologi ke dalam alat pembangunan ontologi. Setelah kelas, atribut dan nilai dikenal pasti maka proses pembangunan ontologi pakar TMK mengguna perisian *Protégé* dilakukan. Melalui perisian ini, hubungan semantik yang wujud dapat dilihat dengan mudah.

Setelah ontologi pakar TMK selesai dibangun, kandungan ontologi ini disah oleh pakar domain. Pengesahan ini penting supaya ontologi pakar TMK yang dibangun berkualiti, membantu dalam pengembangan pengetahuan pengguna dan meningkat kualiti hasil carian pada prototaip sistem.

#### PENGESAHAN KANDUNGAN ONTOLOGI

Pengesahan ontologi kepakaran TMK dilaksana dengan menemu bual beberapa pakar domain yang terdiri daripada pensyarah di FTSM, UKM. Pakar perlu menilai dan memberi komen dari segi skop kajian bidang kepakaran TMK. Pakar domain terdiri daripada ketua dan ahli kumpulan penyelidikan bidang TMK di FTSM. Penerangan diberi kepada pakar domain tentang ontologi kepakaran TMK, peranan ontologi dalam capaian maklumat.

Temu bual dilakukan mengguna kaedah *think aloud* bagi menilai kandungan ontologi. Kaedah ini dipilih kerana melalui perbincangan lanjut dengan pakar domain maka setiap kekurangan dalam ontologi kepakaran TMK dapat dikenal pasti dan seterusnya menambah baik ontologi yang dibangun. Melalui proses ini konsep penting yang perlu ada dalam ontologi kepakaran TMK dapat diketahui.

#### KESIMPULAN

Jika penyelidik terdahulu bergantung kepada pencirian seseorang pakar tanpa melihat dari aspek pengayaan maklumat pakar tersebut, kajian ini melibatkan pakar domain kerana dapat membantu dalam proses memperkaya maklumat dengan merujuk sumber yang pelbagai seperti taksonomi ACM dan tesaurus. Gabungan sumber ini mewujudkan satu kerangka yang merangkumi pencirian pakar dan peluasan hubungan antara sumber yang dipercayai. Ontologi pakar TMK



yang dihasilkan dapat membantu pengguna mendapat padanan pakar yang relevan dan proses carian pakar menjadi cepat dan terfokus kepada bidang carian.

Kerangka pembangunan ontologi kajian ini menjadi titik permulaan kepada kajian yang lain. Misalnya, ontologi yang dibangun boleh diterap ke dalam sebuah sistem pencarian pakar. Penyelidikan juga boleh dijalankan bagi mengukur ketepatan hasil carian pakar TMK yang mengaplikasikan ontologi pakar TMK selain daripada boleh mengaplikasikan kerangka pembangunan ontologi dalam bidang lain seperti kejuruteraan, perubatan dan perniagaan.

Sistem pencarian pakar yang mengaplikasikan ontologi pakar TMK memberi manfaat kepada banyak organisasi seperti Kementerian Pendidikan, Kementerian Sains, Teknologi dan Inovasi, dan universiti sama ada pada peringkat tempatan mahupun luar negara. Pencarian maklumat tentang seseorang pakar perlu bagi memperoleh padanan bidang kepakaran dengan pakar sebelum jalinan kerjasama dalam projek penyelidikan dapat diwujudkan.

## RUJUKAN

- Boyce, S. & Pahl, C. 2007. Developing Domain Ontologies for Course Content. *Educational Technology & Society*, 10(3): 275-288.
- Carvalho, R.M., Berrios, D.C., Wolfe, S.R. & Williams, J. 2005. Ontology Development and Evolution in the Accident Investigation Domain. Paper Presented at IEEE Aerospace Conference, 5-12 December, Montana.
- Chen, L., Shadbolt, N. R., Goble, C., Tao, F. Cox S. J., Puleston, C. & Smart, P. R. 2003. Towards a Knowledge-Based Approach to Semantic Service Composition. *The Semantic Web*, 2870: 319-334.
- Daconta, M. C., Obrst, L. J., & Smith, K. T. 2003. *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge*. Indianapolis: John Wiley & Sons.
- Ensan, F. & Du, W. 2007. Towards Domain-Centric Ontology Development and Maintenance Frameworks. *Proceedings of the 19<sup>th</sup> International Conference on Software Engineering & Knowledge Engineering*. Skokie: Knowledge Systems Institute Graduate School, 622-627.
- Fonou Dombou, J.V. & Huisman, M. 2011. Combining Ontology Development Methodologies and Semantic Web Platforms for E-government Domain Ontology Development. *International Journal of Web & Semantic Technology* 2(2): 12-25.
- Jing, Y. & Croft, W. B. 1994. An association thesaurus for information retrieval. *Proceedings of the 4<sup>th</sup> International Conference on Computer Assisted Information Retrieval*. New York: The Rockefeller University, 146-160.
- Juhana Salim, Mohd. Shahizan Othman & Leng Ten Moi. 2007. *Organisasi Maklumat*. Skudai: Penerbit UTM.
- Liu, P., Liu, K. & Liu, J. 2007. Ontology-based Expertise Matching System within Academia. *Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing*. Shanghai: IEEE, 5431-5434.
- McDonald, D. W. 2001. Evaluating Expertise Recommendations. *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*. Colorado: ACM New York, 214-223.
- Mirkin B, Nascimento S. & Pereira, L.M. 2007. ACM classification can be used for representing research organizations. Technical report written for Center for Discrete Mathematics and Theoretical Computer Science, Rutgers University.
- The 1998 ACM Computing Classification System. <http://www.acm.org/about/class/ccs98-html> [12 May 2014].
- Omoronyia, I., Sindre, G., Stålhane, T., Biffi, S., Moser, T. & Sunindyo, W. 2010. A Domain Ontology Building Process for Guiding Requirements Elicitation. *Proceedings of the 16<sup>th</sup> International Working Conference*. Berlin: Springer Berlin Heidelberg, 188-202.
- Punnarut, P. & Sriharee, G. 2010. A researcher expertise search system using ontology-based data mining. *Proceedings of the 7<sup>th</sup> Asia-Pacific Conference on Conceptual Modelling*. Brisbane: Australian Computer Society, Inc, 71-78.

- Saeed, H. & Chaudhry, A. S. 2002. Using Dewey decimal classification scheme (DDC) for building taxonomies for knowledge organisation. *Journal of Documentation*. 58(5): 575-583.
- Schult, R. & Spiliopoulou, M. 2006. Discovering emerging topics in unlabelled text collections. *Proceedings of the 10<sup>th</sup> East European conference on Advances in Databases and Information Systems*. Berlin: Springer Berlin Heidelberg, 353-366.
- Simperl, E.S.B. & Mochol, M. 2006. Cost Estimation for Ontology Development. *Proceedings of the 9<sup>th</sup> Business Information Systems*. New York: Springer Science & Business Media, 3-16.
- Siti Farhana M. Hashim, Juhana Salim & Shahrul Azman M. Noah. 2014. Domain Ontology Development Framework: The Thesaurus, Reference Source, Knowledge Engineer and Domain Expert Approach. *Asia-Pacific Journal of Information Technology and Multimedia*, 3(2): 71-92.
- Stefanov, K. & Yordanova, K. 2003. Computing Ontology Creation. Proceedings of the International Congress of Mathematical Society of South-East Europe. Bulgaria: Bulgarian academy of sciences, Institute of mathematics and informatics.
- Thesaurus Web IEEE. <http://www.ieee.org/documents/pdfieethes05nov09.pdf> [8 March 2015].
- Velardi, P., Fabriani, P., & Missikoff, M. 2001. Using text processing techniques to automatically enrich a domain ontology. *Proceedings of the International Conference on Formal Ontology in Information Systems*. New York: ACM New York, 270-284.
- Xing, X., Li, R. & Liu, K.Y. 2009. Building Ontology Base on Thesaurus. *Proceedings of the 2<sup>nd</sup> International Conference on Biomedical Engineering and Informatics*, Tianjin: IEEE, 1-4.

Akmal Aris,  
Junaidah Mohamed Kassim,  
Hana Yasmeh Ishak,  
Juhana Salim,  
Shahrul Azman Mohd Noah  
Fakulti Teknologi & Sains Maklumat  
Universiti Kebangsaan Malaysia

Received: 3 April 2015  
Accepted: 1 Jun 2015  
Published: 1 Oktober 2015