

Penghasilan Nombor Pengkelasan Secara Automatik

MOHD. SHAHIZAN OTHMAN, ZARINA SHUKUR,
JUHANA SALIM & MASRI AYOB

ABSTRAK

Dalam sesebuah sistem pengurusan sumber maklumat perpustakaan atau pusat sumber, antara proses penting yang terlibat adalah pengkatalogan dan pengkelasan sumber-sumber maklumat. Kedua-kedua proses ini membolehkan data bagi setiap sumber maklumat dijana, direkod bersama dengan nombor panggilan yang diberi, bagi membolehkan sumber maklumat dikategori, di susun dan dicapai. Nombor panggilan merupakan nombor unik bagi membezakan satu sumber maklumat daripada yang lain dan ia terdiri daripada nombor klasifikasi, yang menentukan bidang sumber berkenaan. Kertas ini membincangkan proses penghasilan nombor pengkelasan secara automatik. Tujuan kajian ini adalah untuk membantu kakitangan sokongan yang tidak mempunyai latarbelakang dalam bidang perpustakaan supaya boleh mengendalikan perpustakaan secara perseorangan. Rekabentuk proses melibatkan konsep capaian maklumat. Implementasi mempertimbangkan bahan dalam bahasa Inggeris dan bahasa Melayu. Ujian ketrampilan proses ini yang telah dilakukan terhadap 200 buah buku, menunjukkan kebolehpercayaan nombor pengkelasan yang dihasilkan.

ABSTRACT

In a library or information center system, among the important processes involved include cataloguing and classification of information sources. These processes enabled data for each information source be generated, recorded along with the call number assigned, to ensure that the sources are categorized, arranged and accessed. The call number is a unique number that differentiates one information source from another and consist of the classification number that determines the subject of the source. This article discusses the processes in generating the classification number by automatic means. The purpose of the research in this aspect is to assist the library clerk, who has no Librarianship background, to manage the library single handedly. The development design processes involved information retrieval concept. The implementation considered information materials in the Malay and English

language. The efficiency and effectiveness of the processes which was conducted on a total of 200 books showed the reliability of the classification number that was generated automatically.

PENGENALAN

Pengkelasan perpustakaan didefinisikan oleh Maltby (1975) sebagai susunan sistematik berdasarkan perkara yang terdapat pada buku dan bahan lain di rak atau pengkatalogan dan entri indeks yang memudahkan pengguna mencari maklumat tersebut. Berdasarkan kepada definisi tersebut, Chan (1994) telah menjelaskan dua objektif pengkelasan perpustakaan iaitu membantu pengguna menentukan lokasi sumber maklumat berdasarkan nombor panggilan dan sumber maklumat yang berkaitan dikumpulkan pada lokasi yang sama. Objektif pertama adalah bermakna kaedah penomboran dan penandaan pada dokumen yang berhubungkait dengan entri pengkatalogan. Objektif kedua pula, memerlukan fungsi kolasi dan sumber maklumat dikumpulkan bersama berdasarkan kriteria yang dipilih. Pengkelasan adalah seperti alatan capaian yang membantu mengenalpasti dan mencapai kumpulan sumber maklumat yang mempunyai perkara yang berkaitan.

Penomboran dan penandaan pada dokumen lazimnya merujuk kepada nombor panggilan yang merupakan simbol komposit yang mengandungi nombor kelas atau nombor item termasuklah data lain seperti tarikh, nombor jilid, yang boleh dijadikan pengenalan sesuatu bahan dan lokasinya di rak (Chan 1994). Nombor panggilan ditentukan berdasarkan skema klasifikasi yang digunakan oleh perpustakaan. Menurut Webster's (1976) pula, nombor panggilan adalah kombinasi abjad untuk mewakili buku di perpustakaan yang boleh menentukan susunannya relatif dengan buku lain di rak.

Dalam memproses sumber maklumat, kelulusan dan kepakaran pustakawan diperlukan kerana pelaksanaan kerja seharian terutamanya pengkatalogan dan pengkelasan merupakan aktiviti teknikal (Savic 1996) yang memerlukan kemahiran asas yang menjadi tunjang profesionalisme kepustakawan. Selain itu, Bocher(1993) menyatakan penyelenggaraan peminjaman dan pengkatalogan adalah berintensifkan buruh tanpa mengambilkira saiz perpustakaan.

Masalah keperluan kepakaran pustakawan bagi menyelenggarakan sumber maklumat turut dialami Pusat Sumber Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia (PSF) yang terpaksa menggunakan kaedah manual di dalam melaksanakan kerja seharian. PSF mempunyai kekangan iaitu kakitangan yang menguruskannya merupakan kakitangan am yang tidak mempunyai latar belakang kepustakawanan. Sistem manual ini menimbulkan pelbagai masalah kepada pembantu pusat sumber dan pengguna seperti yang dinyatakan oleh Adam et al.(1996) antaranya:

- a. Kesukaran dalam menyimpan, mengesan dan mencapai sumber maklumat kerana ketiadaan kemahiran mengkatalog dan mengkelaskan bahan.
- b. Proses penempahan dan maklumat tentang status sesebuah buku adalah tidak berkesan kerana pengguna terpaksa ke pusat sumber untuk menempah atau meminjam sumber maklumat tersebut.
- c. Sukar mengawal pergerakan aliran masuk dan keluar sumber maklumat.
- d. Laporan sukar atau tidak dapat dikeluarkan secara automatik.

Salah satu cara bagi mengatasi masalah yang seperti yang telah dibincangkan di atas, adalah dengan mengimplimentasikan automasi dengan membeli sistem siap guna. Sistem siap guna didefinisikan sebagai sebuah sistem yang lengkap dan boleh digunakan tanpa tambahan atau pengubahsuaian apabila diserahkan kepada pelanggan (Zoraini Wati Abas *et. al.* 1994). Walau bagaimanapun, faktor kos yang tinggi serta ketidaksesuaian pakej dalam persekitaran pusat sumber menyukarkan pengimplementasian sistem siap guna (Juhana Salim *et al.* 2000). Menurut Swan (1996), perbelanjaan untuk membeli dan menyelenggarakan sistem siap guna adalah antara RM 50,000 hingga RM500,000. Ketidakesuaian pakej pula disebabkan oleh kesukaran membuat pengubahsuaian mengikut keperluan semasa (Cohn *et al.* 1997)

JADUAL 1. Kadar harga perkhidmatan yang dikenakan oleh pembekal

Bil.	Perkara	Harga
1.	Pengkatalogan asal (original cataloging)	US\$35 per tajuk
2.	Pengkatalogan menyalin (copy cataloging)	US\$27 per tajuk
3.	Pemberian nombor panggilan atau subjek utama	US\$1.90 per item
4.	Percetakan kod palang	US\$0.20 per item
5.	Percetakan rekod MARC	US\$0.25 per muka surat
6.	Penulisan nombor panggilan pada sumber maklumat	US\$0.15 per item

Sumber: Mei, T. 1998. Outsourcing CJK Cataloging: Implementation Plan.

Selain menggunakan sistem siap guna, pusat sumber boleh menggunakan perkhidmatan yang disediakan oleh pembekal. Kaedah ini mampu menjimatkan kos pembelian peralatan untuk tujuan pengautomasian. Namun, kos perkhidmatan yang disediakan adalah terlalu mahal untuk jangka waktu panjang kerana kosnya berdasarkan perkhidmatan yang diberikan pada sumber maklumat (Mei 1998). Jadual 1 menunjukkan kadar harga perkhidmatan yang dikenakan oleh pembekal.

Kajian ini membincangkan pembangunan sistem pengurusan sumber maklumat yang memudahkan pembantu pusat sumber melaksanakan tugas pengurusan sumber maklumat yang melibatkan pengkatalogan dan pengkelasan

secara persendirian. Kajian ini menumpukan kepada cara untuk menyelesaikan masalah utama iaitu untuk mengkatalog sumber maklumat yang terdapat di pusat sumber. Bagi mengkatalog sumber maklumat di PSF, para penyelidik telah menggunakan sistem pengkatalogan AACR dan skema pengkelasan *Library of Congress (LC)* serta menjanakan sistem pengkelasan yang unik bagi mengkelaskan sumber seperti tesis, kertas-kertas persidangan dan laporan teknikal. Kertas ini hanya fokus kepada proses pengkelasan buku sahaja.

MASALAH PENGKATALOGAN DAN PENKELASAN

Menurut Juhana Salim *et al.* (2000), pembangunan satu katalog merupakan suatu usaha yang kompleks dan hanya mampu dilaksanakan oleh pustakawan yang faham tentang peraturan dalam pengkatalogan dan pengkelasan. Tambahan pula, pengguna sering tidak faham menggunakan katalog tradisional dan persoalan selalu timbul terhadap keberkesanan sistem katalog tradisional yang dianggap merumitkan pengurusan sistem maklumat.

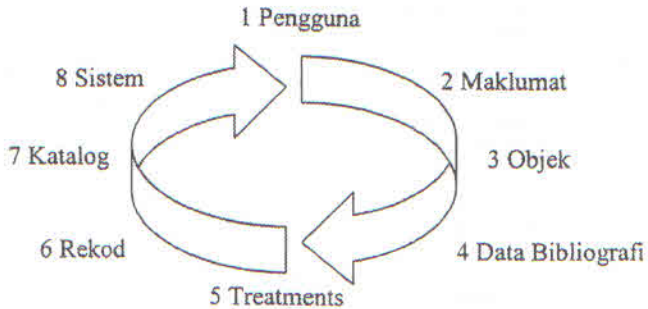
Adalah tidak boleh dinafikan bahawa kompleksiti dalam pengurusan sumber maklumat yang melibatkan pengkatalogan dan pengkelasan, masih dianggap satu keperluan bagi mengelakkan masalah pengurusan sumber maklumat yang akan timbul sekiranya tidak menggunakan cara yang sistematik dalam pengurusan sumber maklumat. Perkembangan dalam automasi boleh memberi satu cara penyelesaian berdasarkan "*machine-generated solution*". Perlu ditekankan bahawa kelemahan sistem pengkatalogan dan pengkelasan yang sedia ada terlalu kompleks strukturnya dan hanya dianggap perlu dalam era pra-automasi. Kompleksiti yang terlibat dalam sistem ini sebenarnya tidak banyak memenuhi kehendak pembantu pusat sumber. Oleh itu kajian ini adalah bukan untuk *consolidate* dengan sistem lama tetapi cuba *capitalise* perkembangan terbaru dalam automasi pusat sumber maklumat/perpustakaan dan komputer.

Dalam pengurusan koleksi perpustakaan yang besar dari segi saiz koleksi, peraturan seperti *Anglo American Cataloguing Rules (AACR2)* dan skema pengkelasan perlu dipatuhi untuk memastikan keseragaman. Hakikatnya, pengurusan PSF yang dipertanggungjawab kepada kakitangan am yang tiada kemahiran mengkatalog dan mengkelaskan sumber maklumat pasti tidak mampu mematuhi peraturan AACR dan sistem pengkelasan LC atau pengkelasan perpuluhan Dewey (DDC). Menyedari masalah di atas, suatu sistem pengkatalogan dan pengkelasan secara automatik perlu dibangunkan bagi membolehkan kakitangan am melaksanakan tugas pemprosesan sumber maklumat dibuat dengan mudah dan cepat.

PENKATALOGAN

Katalog didefinisikan oleh *Anglo-American Cataloguing Rules* (1998) sebagai senarai bahan yang terdapat di dalam koleksi, perpustakaan atau kumpulan perpustakaan yang disusun berdasarkan peraturan-peraturan tertentu.

Pada tahun 1876, Cutter (1953) telah memberikan objektif sistem pengkatalogan perpustakaan yang perlu dipenuhi. Ianya terdiri daripada tiga fungsi iaitu mengetahui item yang dicari, kategori carian dan pemilihan item.



RAJAH 1. Kitaran Lengkap Proses Pengkatalogan

Bagi mencapai objektif sistem pengkatalogan, menurut Jeng (1997) pengkatalog perlulah mengetahui kitaran lengkap proses pengkatalogan (Rajah 1). Kitaran lengkap proses pengkatalogan bermula dengan pengguna yang ingin mencapai maklumat bersama objek maklumat. Tugas pengkatalog adalah untuk memibna hubungan daripada maklumat diperlukan kepada bahagian teks dalam maklumat objek yang memenuhi keperluan. Proses ini memerlukan pengkatalog memahami rekabentuk sistem dan penilaian untuk mengintegrasikan kueri maklum balas pengguna bagi mendapatkan hasil carian.

Pengkatalogan merupakan persediaan maklumat bibliografi untuk rekod katalog (Mortimer 1998). Pengkatalogan terdiri daripada tiga bahagian iaitu (Rajah 2)

a. Pengkelasan

Pengkelasan menentukan nombor pengkelasan untuk item. Pengkelasan boleh dilakukan dengan menggunakan skema pengkelasan LC atau DCC yang menentukan lokasi item pada koleksi.

b. Pengkatalogan deskriptif

Satu proses pengkatalogan yang berkaitan dengan pengenalan dan huraian muka surat judul (kenyataan penanggungjawaban, edisi, butiran khusus bahan dan perihai penerbitan siri) dan kolasi (huraian fizikal bahan), nota dan nombor piawaian buku antarabangsa (ISBN).

c. Pengkatalogan perkara

Pengkatalogan perkara menentukan tajuk perkara dengan menggunakan *Library of Congress Subject Headings* atau *Sears List*.

CALL NUMBER	Reference collection RF 020.3 H323-8	} Pengkelasan
TITLE	Harrod's Librarian's glossary : 9000 terms used in information management, library science, publishing, the book trades and archive management / compiled by Ray Prytherch.	
EDITION	8th ed.	} Pengkatalogan deskriptif
PUBLISHER	Aldershot, Hants, England : Gower ; Brookfield, Vt., USA : Ashgate, c1995.	
DESCRIPT'N	xiii, 692 p. ; 25 cm.	
AUTHOR(S)	Prytherch, Raymond John	
SUBJECT	1. Library science -- Dictionaries 2. Information science -- Dictionaries 3. Publishers and publishing -- Dictionaries 4. Book industries and trade -- Dictionaries 5. Archives -- Dictionaries 6. Bibliography -- Dictionaries	} Pengkatalogan perkara

RAJAH 2. Rekod Pengkatalogan

Pengorganisasian sumber maklumat tidak lengkap sekiranya sumber maklumat yang dikatalog tidak diberi nombor pengkelasan. Proses mengenal pasti nombor pengkelasan yang perlu diberi bagi sesuatu sumber dikenali sebagai pengkelasan. Seterusnya, kertas ini membincangkan bahagian pengkelasan. Ini kerana sekiranya pengkelasan dokumen dilakukan secara manual, ianya memakan masa dan memerlukan kemahiran dan kecekapan pustakawan (Siti Hasnah Simanjuntak 1984).

PENGKELASAN

Pada ketika ini, terdapat dua sistem pengkelasan yang utama dan digunakan oleh perpustakaan sekarang. Sistem tersebut ialah skema pengkelasan Library of Congress (LC) dan Sistem Pengkelasan Perpuluhan Dewey (DDC). Kedua-dua skema pengkelasan ini berkembang mengikut masa dan menggambarkan kepada kemuncak tahap intelektual am pada era kini (Chan 1994).

Terdapat langkah yang sistematis untuk mengelaskan dokumen. Pustakawan perlu mengikuti sepuluh langkah untuk menghasilkan nombor panggilan untuk buku. Sepuluh langkah untuk pengelasan dokumen menurut Dittman dan Hardy (2000) ialah:

1. Periksa sumber maklumat untuk menentukan perkara.
2. Lihat pada tajuk perkara.
3. Tulis kata kunci yang mewakili perkara.
4. Pilih Jadual Pengelasan LC yang relevan.
5. Gunakan pengindeksan pada jadual sebagai titik permulaan melakukan pencarian nombor pengelasan.
6. Lihat pada kandungan jadual LC dan pilih nombor pengelasan yang sesuai.
7. Tentukan nombor untuk entri utama.
8. Tambahkan tarikh penerbitan.
9. Semak bahan di rak dan buat pengubahsuaian sekiranya perlu.
10. Sumber maklumat sedia untuk proses pelabelan dan disusun di rak.

Langkah 1 hingga 3 merupakan langkah untuk menentukan perkara. Sekiranya pustakawan ingin mendapatkan maklumat berkenaan dengan pengelasan yang tepat, pemeriksaan atas perkara berikut diperlukan iaitu:

1. Tajuk - mungkin atau tidak membantu.
2. Tajuk lain - selalunya membantu.
3. Jadual kandungan - digunakan untuk menentukan topik utama.
4. Kata pengantar, prakata, pengenalan - selalunya nyatakan berkenaan apa yang penulis tekankan dan subjek yang diliputi.
5. Kulit buku - memberikan maklumat berkenaan dengan pengarang dan ringkasan.
6. Teks - digunakan untuk mengesahkan idea berkenaan dengan perkara.

Hasil akhir dalam proses pengkatalogan bahan adalah menentukan nombor panggilan untuk sumber maklumat. Nombor panggilan sebenarnya merupakan gabungan nombor pengelasan (yang dipilih dari skema pengelasan LC atau DDC) dengan nombor cutter pengarang.

Sehingga kini, terdapat beberapa perisian untuk membantu pengelasan seperti Perisian *Cutt-x*, Perisian *FastCat* dan perisian *The Cataloging Calculator*. *Cutt-x* telah dibangunkan oleh Savic(1996) untuk penentuan nombor Cutter secara automatik. Manakala Perisian *FastCat* telah dibangunkan di *University of Nevada, Las Vegas* (Downs 1998). Ianya menumpukan kepada capaian rekod MARC yang terdapat pada katalog atas talian. *The Cataloging Calculator* pula dibangunkan di *Oregon State University* oleh Banerjee (2001). Perisian ini digunakan bertujuan untuk menjanakan nombor Cutter LC.

Daripada ketiga-tiga perisian tersebut, tiada yang menjurus ke bahagian pengautomasian nombor pengkelasan. Kertas ini mencadangkan proses pengkelasan secara berautomasi dengan menggunakan konsep capaian maklumat.

KONSEP CAPAIAN MAKLUMAT

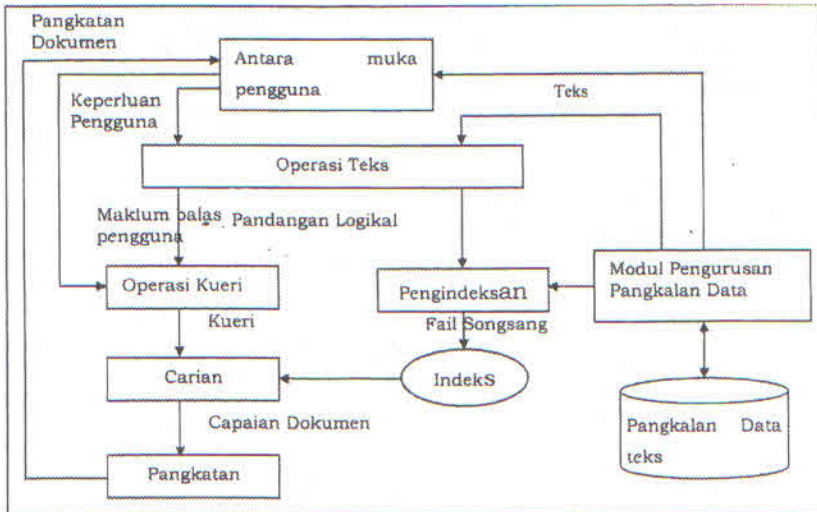
Capaian maklumat merupakan proses pertama yang di lalui sebelum nombor panggilan dapat dihasilkan. Seperti yang telah dinyatakan sebelum ini, nombor panggilan terhasil daripada gabungan nombor pengkelasan dan nombor Cutter. Konsep capaian maklumat digunakan untuk mendapatkan susunan pangkatan yang relevan berdasarkan skema pengkelasan LC bagi menghasilkan nombor pengkelasan sumber maklumat tersebut. Perbincangan lanjut berkaitan penghasilan nombor Cutter pula boleh didapati daripada tesis Sarjana oleh Mohd. Shahizan Othman (2001).

Konsep capaian maklumat ialah berhubungan dengan persembahan, storan, organisasi dan capaian kepada item maklumat (Baeza Yates dan Ribeiro-Neto 1999). Persembahan dan organisasi item maklumat memberi kesenangan kepada pengguna dari segi mencapai maklumat yang dikehendaki. Bagaimanapun, kriteria untuk memenuhi keperluan maklumat pengguna bukan merupakan perkara yang mudah. Pengguna perlu menterjemahkan maklumat yang diperlukan kepada bentuk kueri dan kemudiannya diproses menggunakan sistem capaian maklumat dan seterusnya pangkatan disenaraikan. Rajah 3 menunjukkan proses capaian maklumat menurut Baeza Yates dan Ribeiro-Neto (1999). Berdasarkan rajah tersebut, secara ringkasnya, proses capaian maklumat melibatkan teks atau kueri yang ingin dicari dan suatu pangkalan data teks. Teks atau kueri tersebut dimasukkan oleh pengguna. Manakala teks dalam pangkalan data pula perlu diindekskan sebelum carian terhadap teks atau kueri yang dimasukkan oleh pengguna tadi boleh dilakukan.

Frakes (1992) telah menghasilkan reka bentuk pandangan fungsian sistem capaian maklumat (Rajah 4). Secara umumnya, dalam rajah tersebut, teks dalam dokumen yang ingin dicari perlu dipecahkan kepada perkataan, kemudian kata henti dibuang, dan seterusnya perkataan terkini itu akan dipangkas. Perkataan yang telah dipangkaskan kemudiannya diberi pemberat. Pemberat merupakan kaedah untuk menentukan tahap kepentingan perkataan di dalam dokumen. Skema pemberat yang digunakan berdasarkan kiraan pemberat $TF*IDF$, di mana:

TF ialah kekerapan perkataan dalam dokumen khusus dan formulanya:

$$TF = \log(\text{frekuensi perkataan } i \text{ dalam dokumen } j) + 1$$
$$\log(\text{frekuensi perkataan unik dalam dokumen } j)$$



RAJAH 3. Pandangan Logikal Proses Capaian Maklumat oleh Baeza Yates dan Ribeiro-Neto (1999)

IDF ialah frekuensi songsangan dokumen perkataan i dan formulanya:

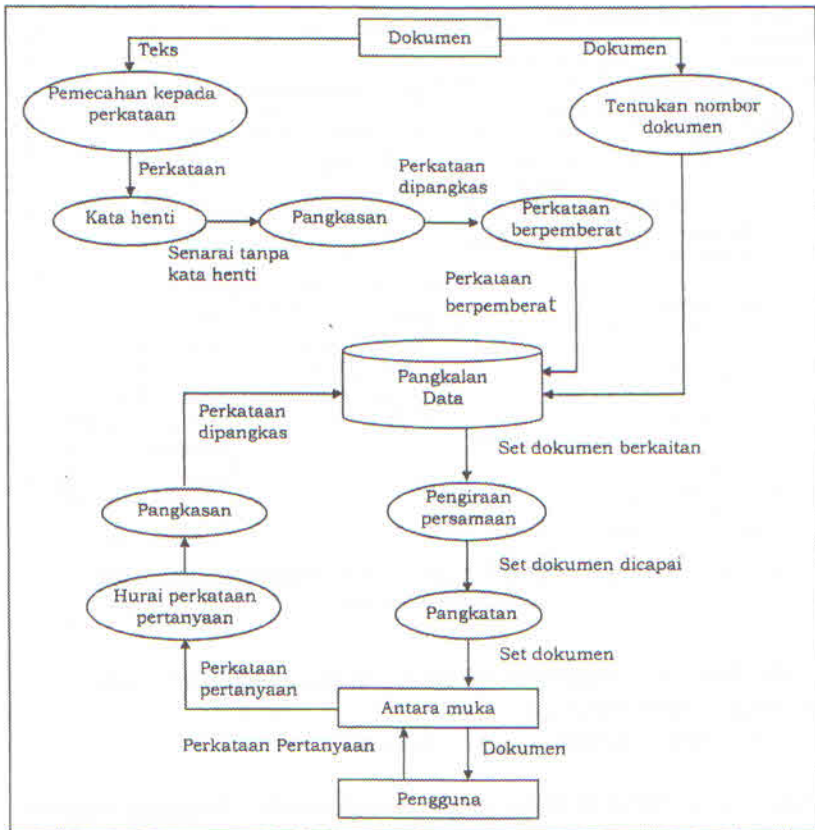
$$IDF = \log_2 \left(\frac{\text{bilangan dokumen}}{\text{bilangan dokumen dengan perkataan } i} \right) + 1$$

Perkataan berpemberat ini disimpan dalam pangkalan data. Manakala perkataan atau kueri yang dimasukkan oleh pengguna turut perlu dipangkas. Kemudian perkataan yang dipangkaskan itu, dicari dalam pangkalan data untuk mendapat pemberatnya bagi proses pangkatan.

Konsep capaian maklumat yang berpangkalan menurut Cutler (1999), mempunyai kelebihan dan kekurangannya. Kelebihan konsep ini adalah dokumen yang dicapai akan disenaraikan mengikut pangkatan keutamaan. Manakala penglibatan proses pengiraan pemberat mengakibatkan ianya lebih kompleks untuk diimplementasi untuk pangkalan data bersaiz besar.

Fungsian sistem capaian maklumat Model Frakes ini merupakan sistem capaian maklumat *Boolean* yang merupakan operasi piawai untuk sistem capaian maklumat. Operator *Boolean* iaitu *AND*, *OR* dan *NOT* digunakan untuk melakukan carian pada pangkalan data setelah kueri dimasukkan oleh pengguna. Model Frakes ini adalah hampir sama dengan proses capaian maklumat oleh Baeza Yates dan Ribeiro-Neto (1999).

Konsep capaian maklumat berdasarkan rekabentuk Model Frakes (1992) dan Baeza Yates dan Ribeiro-Neto (1999) hanya digunakan untuk menghasilkan pangkatan bagi skema pengkelasan yang relevan.



RAJAH 4. Pandangan Fungsian Sistem Capaian Maklumat oleh Frakes (1992)

REKA BENTUK MODEL KONSEPSI PENGKATALOGAN DAN PENGKELASAN

Penghasilan reka bentuk kajian ini adalah berdasarkan konsep capaian maklumat yang dikemukakan oleh Baeza-Yates (1999) dan Frakes (1992). Kedua-dua konsep ini (seperti yang telah dibincangkan dalam Bahagian 5.0) diintegrasikan untuk menghasilkan reka bentuk model konsepsi pengkelasan.

Konsep capaian maklumat digunakan dari proses kemasukan data sehinggalah kepada penghasilan pangkatan yang relevan berkenaan sumber maklumat tersebut berdasarkan skema pengkelasan LC. Setelah nombor pengkelasan diberikan, nombor panggilan dihasilkan daripada gabungan nombor skema pengkelasan LC dan nombor Cutter.

IMPLIMENTASI PROSES PENGKELASAN SUMBER MAKLUMAT

Modul pemprosesan menggunakan kaedah capaian maklumat untuk menghasilkan paparan dokumen pengkelasan mengikut susunan pangkatan. Modul pemprosesan bermula dengan kemasukan sumber maklumat oleh pembantu pusat sumber. Tajuk sumber maklumat diperlukan untuk membolehkan penentuan nombor pengkelasan sumber maklumat. Rajah 5 menunjukkan proses keseluruhan modul pemprosesan.

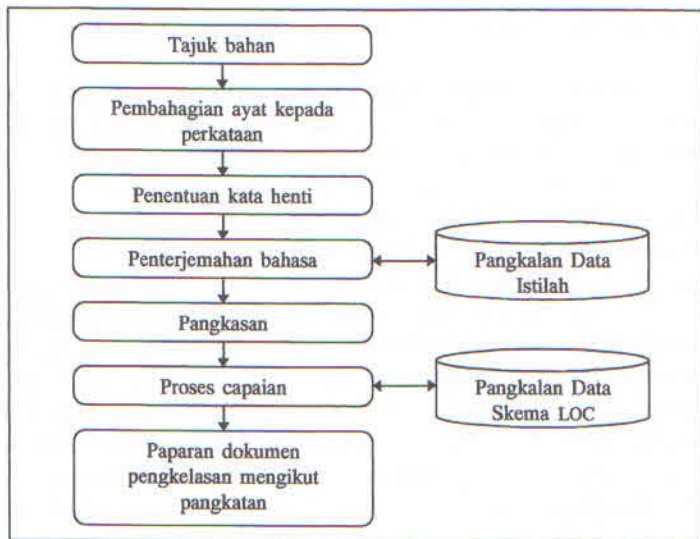
Modul pemprosesan membaca input tajuk sumber maklumat satu persatu. Ayat yang dibaca ini kemudiannya dibahagikan kepada perkataan-perkataan. Setiap perkataan dikenal pasti melalui penyemakan satu persatu huruf sehingga ruang kosong ditemui. Sekiranya semasa huruf dibaca dan tanda ‘-’ ditemui, maka perkataan sebelum tanda ‘-’ diambil sebagai satu perkataan dan huruf seterusnya akan dibaca sehingga ruang kosong. Jika perkataan ini sama dengan sebelumnya, perkataan tersebut diabaikan dan perkataan seterusnya dianggap sebagai satu perkataan baru. Contohnya adalah seperti berikut:

Tajuk Asal : Pengauditan, kawalan dan keselamatan komputer
Perkataan yang wujud : “Pengauditan”, “kawalan”, “dan”, “keselamatan”, “komputer”

Setelah setiap ayat dibahagikan kepada perkataan, proses penentuan kata henti akan dilaksanakan. Proses ini mengenal pasti sama ada perkataan yang wujud merupakan kata henti ataupun sebaliknya. Perkataan tersebut dibandingkan dengan senarai kata henti Bahasa Melayu ataupun Bahasa Inggeris. Jika perkataan tersebut sama dengan perkataan dalam senarai kata henti, perkataan tersebut akan dibuang daripada meneruskan proses selanjutnya. Contoh perkataan kata henti adalah “**untuk**”, “**itu**”, “**yang**”, “**dan**”, “**atau**” dan “**the**”, “**is**”, “**or**”, “**and**” dalam Bahasa Inggeris (Rujuk lampiran A dan B). Bagi contoh di atas:

Perkataan yang wujud : “Pengauditan”, “kawalan”, “dan”, “keselamatan”, “komputer”
Selepas penentuan kata henti: “Pengauditan”, “kawalan”, “keselamatan”, “komputer”

Jika sumber maklumat adalah di dalam Bahasa Melayu, proses penterjemahan bahasa diperlukan. Pangkalan data Istilah digunakan untuk membuat terjemahan dari perkataan Bahasa Melayu dan Bahasa Inggeris. Perkataan yang tiada dalam senarai kata henti dipadankan dengan perkataan yang terdapat di dalam pangkalan data Istilah. Rajah 6 menunjukkan format Pangkalan Data Istilah.



RAJAH 5. Modul Pemrosesan

Bahasa Melayu	Bahasa Inggeris
Aalen Abadi	Aalenian Conservative
↓	↓
Zuriat Zwiterion	Offspring Zwiterion

RAJAH 6. Format Pangkalan Data Istilah

Sekiranya perkataan Bahasa Melayu adalah “*Aalen*” proses pepadanan dijalankan dan mendapati perkataan penterjemahan di dalam Bahasa Inggeris adalah “*Aalenian*”. Penterjemahan bahasa diperlukan kerana maklumat skema pengkelasan LC adalah dalam Bahasa Inggeris. Sistem menterjemahkan perkataan tersebut dan kemudiannya proses pangkasan dilaksanakan. Bagi contoh yang sebelum ini:

Perkataan yang wujud : “Pengauditan”, “kawalan”, “keselamatan”, “komputer”

Selepas terjemahan: “audit”, “control”, “security”, “computer”

Pangkasan merupakan proses yang membuang semua imbuhan yang terdapat dalam sesuatu perkataan. Proses ini dilakukan untuk meningkatkan keupayaan capaian dan menjimatkan ruangan storan. Penyelidik menggunakan

algoritma pangkasan Porter (1980) untuk Bahasa Inggeris. Algoritma ini bertujuan untuk mendapatkan kata akar perkataan Bahasa Inggeris. Algoritma ini telah digunakan dengan meluas semenjak 20 tahun yang lalu. Proses pangkasan dilakukan seperti perkataan *computers*, *computing*, *computerized* dan *computerization* menghasilkan kata akar iaitu *comput*. Bagi contoh yang sebelum ini:

Perkataan yang wujud : “audit”, “control”, “security”, “compute”
Selepas pangkasan: “audit”, “control”, “secure”, “comput”

Setelah itu, proses capaian kepada perkataan indeks yang terdapat pada fail songsang (*inverted file*) dilakukan. Fail songsang terdiri daripada fail dokumen dan fail indeks. Fail dokumen dalam kajian ini merupakan dokumen yang terdapat di pangkalan data skema pengkelasan LC. Oleh kerana tujuan asal kajian ini adalah untuk kegunaan pusat sumber khusus iaitu PSF yang mempunyai sebahagian besar bahan media di dalam kategori Matematik dan Sains Komputer (QA) dan Sains Perpustakaan (Z), maka hanya skema pengkelasan LC, kelas QA dan Z sahaja yang diimplemen. Lebih khusus lagi, fail dokumen ini hanya terdiri daripada nombor panggilan dan keterangannya.

Fail indeks pula terhasil daripada proses pengindeksan terhadap fail dokumen. Ianya mengandungi kata kunci, bilangan kekerapan dan pautan. Setiap dokumen ditandakan dengan senarai kata kunci, (kata kunci merujuk kepada perkataan yang dipangkaskan kecuali kata henti) dengan setiap satunya mempunyai bilangan kekerapan kewujudan perkataan tersebut dalam dokumen itu. Setiap kata kunci ini mempunyai hubungan dengan dokumen-dokumen lain yang mengandungi kata kunci yang sama.

Proses capaian menggunakan kaedah cincangan untuk memastikan proses yang lebih pantas (Belal 1994). Proses capaian dilakukan ke atas fail dokumen yang terdapat pada pangkalan data skema LC. Fail dokumen kemudiannya disusun berdasarkan hasil operasi Boolean serta nilai bilangan kekerapan dan dipaparkan untuk membolehkan pembantu pusat sumber memilih pengkelasan untuk sumber maklumat. Rajah 7 menunjukkan konsep fail songsang untuk mencapai dokumen di pangkalan data skema LC.

ANALISA KETRAMPILAN PENGOTOMASIAN PROSES PENKELASAN BUKU

Penyelidik telah menjalankan pengujian ke atas 200 buah buku yang dipilih secara rawak. Maklumat berkenaan dengan buku ini diambil daripada Perpustakaan Tun Sri Lanang, Universiti Kebangsaan Malaysia dan Perpustakaan Sultanah Zanariah, Universiti Teknologi Malaysia. Buku-buku ini terdiri daripada 183 buah buku Bahasa Inggeris dan 17 buah buku Bahasa Melayu. Pengujian ini melibatkan skema pengkelasan LC iaitu QA sebanyak 167 buah dan Z sebanyak 33 buah.

- a. Perisian tidak dapat mengesan nombor panggilan buku yang diletak di bawah pengkelasan karya umum kerana nombor panggilan buku tersebut tidak terdapat dalam pangkalan data sumber maklumat.
- b. Kegagalan memadankan tajuk dengan pangkalan data LC disebabkan oleh penggunaan perkataan yang berlainan, walaupun mempunyai maksud yang sama.
- c. Perisian mengekalkan perkataan dalam Bahasa Melayu dan dianggapnya sebagai nama khas jika didapati perkataan tersebut tiada dalam pangkalan data Istilah. Ini dilakukan semasa membuat penterjemahan tajuk buku Bahasa Melayu kepada Bahasa Inggeris.
- d. Masalah kesalahan penterjemahan daripada Bahasa Melayu kepada Bahasa Inggeris yang boleh diatasi dengan menggunakan tesaurus.

ANALISIS UNTUK BUKU YANG MEMPUNYAI PANGKATAN

Pangkalan akan diberikan berdasarkan kepada tahap ketepatan dan kerelevanan yang tinggi di antara tajuk buku dengan pangkalan data LC. Sekiranya buku tersebut mempunyai ketepatan dan kerelevanan yang tinggi, pangkatan pertama akan diberikan. Nilai pangkatan akan bertambah sekiranya keadaan sebaliknya berlaku. Jadual 2 ialah contoh berkenaan dengan pelaksanaan pangkatan kedua untuk buku.

JADUAL 2. Pelaksanaan pangkatan kedua untuk buku

Perkara	Keterangan
No. panggil	QA76.9.A93
Tajuk	Computer audit, control and security
Penerangan	Pangkalan pertama adalah QA76.9.A25 iaitu <i>Access Control and Computer Security</i> . Maklumat skema LOC untuk QA76.9.A93 ialah <i>Auditing</i> . Walaupun perisian berjaya menjumpai nombor panggilan untuk tajuk buku tersebut, tetapi pangkatan pertama iaitu QA76.9.A25 adalah lebih tepat. Ketepatan ini disebabkan terdapat perkataan "computer", "control", "security" berbanding skema LC iaitu "auditing"

Hasil pengujian terhadap buku yang mempunyai pangkatan mendapati bahawa, perisian melaksanakan pengujian berdasarkan ketepatan tajuk dengan pangkalan data LC. Buku yang mempunyai nilai pangkatan selain daripada satu tidak boleh dianggap salah kerana terdapat tajuk buku yang tidak menggambarkan kandungannya. Ini adalah kerana perisian hanya melakukan pengujian ke atas tajuk buku tersebut. Walaupun pangkatan buku adalah besar, nombor panggilan tersebut boleh diterima pakai dengan syarat pembantu pusat sumber perlu memastikan maklumat sebenar berdasarkan maklumat bibliografik atau abstrak buku.

Terdapat nombor panggilan yang mempunyai lebih daripada satu lokasi. Contohnya bagi buku yang bertajuk "Pascal", ia terdapat pada lokasi QA76.73.P2 dan Z678.93.P35. Kedua-dua nombor panggilan ini adalah betul dan untuk memastikan lokasi yang sebenar, pembantu pusat sumber perlu menggunakan cara manual yang dibincangkan dalam Bahagian 4.0.

KESIMPULAN

Kertas ini membincangkan keperluan kepada pengautomasian proses pengkelasan, begitu juga proses bagaimana nombor pengkelasan dihasilkan secara berautomasi. Seterusnya analisa ketrampilan proses penghasilan nombor pengkelasan secara berautomasi telah dibentangkan.

Daripada hasil analisis, dapat disimpulkan bahawa sistem ini berkeupayaan menghasilkan nombor pengkelasan secara automatik. Ketepatan sebanyak 66% telah menghasilkan nombor pengkelasan dan ini menunjukkan sistem ini boleh diimplementasi bagi memudahkan pengorganisasian PSF. Manakala 7% bagi pangkatan kedua dan ketiga tidak membimbangkan kerana didapati ianya masih dalam perkara maklumat yang boleh diterima. Seterusnya 20.5% bagi pangkatan antara 4 dan 40 adalah perbezaan yang merujuk kepada lokasi. Masalah utama iaitu menentukan nombor panggilan sumber maklumat yang berdasarkan kepada tajuk boleh diatasi dengan menambah abstrak bahan semasa carian nombor panggilan dilakukan. Selain itu, pembantu pusat sumber perlu melaksanakan proses yang sama seperti menentukan nombor panggilan untuk karya umum.

Hasil kajian ini dapat memberi manfaat kepada pusat sumber unik khasnya PSF dan implementasinya boleh disesuaikan untuk penggunaan pusat sumber khas yang lain dengan cara membina pangkalan data pengkelasan LC yang bersesuaian kepada koleksi sesebuah pusat sumber.

RUJUKAN

- Adam, N.R., Holowczak, R. Halem, M., & Lal, N. 1996. *Digital Library Task Force*. IEEE Journal. 8:89-91
- Anglo-American Cataloguing Rules. Ed. ke-2., 1988. Chicago: American Library Association.
- Baeza-Yates, R., & Riebiero-Neto, B. 1999. *Modern Information Retrieval*. Essex: Addison Wesley Longman Limited.
- Banerjee, K. 2001. *The Cataloging Calculator*. (atas talian) <http://ucs.orst.edu/~banerjek/cutter.html> (10 Mac 2001)
- Belal Abu Ata. 1994. *Design and Implementation of multilingual document retrieval system*. Tesis Sarjana. Universiti Kebangsaan Malaysia.
- Bocher, B. 1993. Small-automated libraries systems. *Computers in Libraries*. 13(2): 26-27.
- Chan, L. M. 1994. *Cataloging and Classification Introduction*. Ed. ke-2. New York: McGraw-Hill.
- Cohn, J.M, Kelsey, A.L, & Files, K.M. 1997. *Planning for automation: a how-to-do it manual for librarians*. Ed. ke-2. New York: Neal-Schuman.

- Cutler, M. 1999. CS533 Information Retrieval (atas talian). http://www.cs.binghamton.edu/~cutler/cs533/lec1_int.html. (21 Mac 2001)
- Cutter C.A. 1953. *Rules for a Dictionary Catalog*. Ed. ke-4. London: The Library Association.
- Dittman, H., & Hardy, J. 2000. *Learn Library of Congress Classification*. Lanham, MD: Scarecrow Press.
- Downs, L. 1998. *Fastcat for Windows and Desert Fastcat for Windows 95*. (atas talian) <http://library.nevada.edu/free/fastcat/> (7 April 2001)
- Frakes, W.B. & Baeza-Yates, R. 1992. *Information Retrieval: Data Structures & Algorithms*. Englewood Cliffs, N.J.: Prentice Hall
- Hayes, R.M. 2000. *Information Storage & Retrieval*. *Encyclopedia Americana Online*. Grolier Inc (atas talian) <http://ea.grolier.com> (18 Nov 2000)
- Jeng, L.H. 1997. *Knowledge, Technology and Research in Cataloging*. *Cataloging & Classification Quarterly* 24:13-127
- Juhana Salim, Zarina Shukur, Syahanim Mohd. Salleh, Rosilah Hassan, Abdullah Mohd. Zin, Mohd. Shahizan Othman, Masri Ayob & Hazrina Harun. (2000). *Rekabentuk Sistem Pengurusan dan Capaian Maklumat, Pusat Sumber Maklumat Pintar, FTSM, UKM*. Kertas Kerja Seminar Kebangsaan Perpustakaan di Malaysia, 23-24 Ogos.
- Maltby, A. 1975. *Sayers' Manual of Classification for Librarians*. Ed. ke-5. London: Andre Deutsch.
- Mei, T. 1998. *Outsourcing CJK Cataloging: Implementation Plan*. (atas talian) <http://www.library.cornell.edu/cts/cjkoutso.htm#appendix1> (1 April 2001).
- Mohd. Shahizan Othman. 2001. *Sistem Pengurusan Pusat Sumber Maklumat: Proses Pengkatalogan dan Pengkelasan Secara Automatik*. Tesis Sarjana. Universiti Kebangsaan Malaysia.
- Mortimer, M. 1998. *Learn Descriptive Cataloguing*. Canberra: DocMatrix Pty Ltd.
- Nicholas, C. 2001. *Overview of Information Retrieval*. (atas talian) <http://www.cs.umbc.edu/~nicholas/courses/491/html/IR.html>. (22 Mac 2001)
- Porter, M. 1980. *An algorithm for sux stripping*. *Program*, Vol. 14(3), pp. 130-137.
- Savic, D. 1996. *CUTT-x: an expert system for automatic assignment of cutter numbers*. *Cataloging and Classification Quarterly* 22(2): 71-87
- Siti Hasnah Simanjuntak. 1994. *Pengkatalogan dan Pengkelasan Bahan-bahan Pustaka*. Shah Alam: Institut Teknologi Mara.
- Webster's Third New International of English Language. 1976. Springfield, Mass: G & C Merriam Company.
- Zoraini Wati Abas, Hoong, N.K., & Seng, G.K. 1994. *Kamus Dwibahasa Komputer*. Kuala Lumpur : Federal Publication Sdn. Bhd.

Mohd. Shahizan Othman
 Fakulti Sains Komputer dan Sistem Maklumat
 Universiti Teknologi Malaysia
 81310 UTM, Skudai, Johor, Malaysia.

Zarina Shukur, Juhana Salim & Masri Ayob
 Fakulti Teknologi dan Sains Maklumat
 Universiti Kebangsaan Malaysia
 43600 UKM, Bangi, Selangor D.E, Malaysia.
 zs@ftsm.ukm.my; js@ftsm.ukm.my