

METADATA, DESKRIPSI SERTA TITIK AKSESNYA DAN INDOMARC

Sulistyo-Basuki
JIP-FSUI

Abstract

The term of metadata often appeared in the literature about database management systems (DBMS) in 1980's. This term is used to note an information characteristics on database. So many definitions about metadata. Metadata is also defined as data of data. Metadata subscribes a source, shows a document location, and gives the summary.

In general, there are three things to make metadata as an information package, i.e: encoding, description of information package and gives access to the description. This paper describes a concept of data relates to the library. The explanation includes a definition of metadata; metadata function; encoding standard, bibliography seconds, surrogate, metadata; creates the contents of surrogate records; design format of metadata; metadata and its standard.

Abstrak

Istilah metadata mulai sering muncul dalam literature tentang database management systems (DBMS) pada tahun 1980 an. Istilah tersebut digunakan untuk menggambarkan informasi yang diperlukan untuk mencatat karakteristik informasi yang terdapat pada pangkalan data. Banyak sumber yang mengartikan istilah metadata. Metadata dapat diartikan sumber, menunjukkan lokasi dokumen, serta memberikan ringkasan yang diperlukan untuk memanfaatkannya. Secara umum ada 3 bagian yang digunakan untuk membuat metadata sebagai sebuah paket informasi, dan penyandian (encoding) pembuatan deskripsi paket informasi, dan penyediaan akses terhadap deskripsi tersebut. Dalam makalah ini diuraikan mengenai konsep data dalam kaitannya dengan perpustakaan. Uraian meliputi definisi metadata; fungsi metadata; standar penyandian (encoding), cantuman bibliografis, surogat, metadata; penciptaan isi cantuman surogat; ancatan terhadap format metadata; serta metadata dan standar metadata.

1. Pengantar

Istilah metadata sudah ada sejak tahun 1960an namun waktu itu belum dikenal di dunia perpustakaan. Istilah tersebut mulai sering muncul dalam literatur tentang *database management systems* (DBMS) pada tahun 1980an. Istilah tersebut digunakan untuk menggambarkan informasi yang diperlukan untuk mencatat karakteristik informasi yang terdapat pada pangkalan data (database). Dalam domain DBMS metadata diberi definisi sebagai data tentang data. Definisi tersebut merupakan dasar bagi definisi yang dibuat kemudian hari. Walaupun definisi metadata tidak menyampingkan data nonelektronik, dalam kenyataannya definisi metadata diterapkan pada data dalam bentuk elektronik.

Karena makalah ini bermaksud menguraikan konsep data dalam kaitannya dengan perpustakaan, maka terlebih dahulu kita menengok perkembangan perpustakaan. Keberadaan perpustakaan sudah berusia lebih dari tiga ribu tahun, dimulai dengan munculnya perpustakaan purba sampai ke perpustakaan elektronik atau perpustakaan digital. Selama perkembangan ribuan tahun itu perpustakaan melaksanakan aktivitas intelektual dalam bentuk menentukan lokasi, mengenali, temu balik dan manipulasi informasi. Selama itu perpustakaan mendeskripsi objek non-elektronik dengan menggunakan per-aturan deskriptif. Objek non-elektronik tersebut di lingkungan perpustakaan dikenal sebagai data bibliografis atau data pengkatalogan yang digunakan untuk mengolah informasi surogat. Tatkala metoda pengorganisasian informasi dari ilmu perpustakaan dan ilmu informasi memasuki lingkungan elektronik, maka istilah metadata yang memiliki konotasi untuk mendeskripsi data elektronik menjadi bagian dari kedua disiplin tersebut. Namun istilah tersebut memperoleh reaksi keras dari dunia pengkatalogan karena mereka menganggap bahwa metadata atau data tentang data tidak lain adalah pengkatalogan atau data bibliografis dalam bentuk nama lain (Larsgaard 1996). Penulis semacam Caplan (1995) dapat

menerima metadata sebagai pengganti istilah pengkatalogan karena metadata merupakan istilah yang bersifat netral tanpa ada makna merendahkan. Berbeda halnya dengan istilah cantuman katalog yang mengandung konotasi negatif bagi berbagai kalangan yang berada di luar ranah (domain) pengkatalogan.

2. Definisi metadata

Berbagai sumber mengatakan bahwa metadata artinya data tentang data (Gritton 1994). Definisi tersebut tidak selalu menimbulkan kemudahan, karena tidak jelas apa yang dimaksudkan dengan data tentang data. Ada yang mengatakan bahwa metadata merupakan dokumentasi tentang dokumen dan objek. Metadata mendeskripsikan sumber, menunjukkan di mana lokasi dokumen serta memberikan ringkasan apa yang diperlukan untuk memanfaatkannya.

Definisi yang diberikan oleh World Wide Web Consortium (1998) menyatakan metadata sebagai mesin yang dapat memahami informasi tentang objek Web serta menyatakan bahwa metadata dapat dikembangkan ke sumber daya elektronik (*electronic resources*) lainnya pada masa depan. Ng et al (1997) memberikan definisi operasional metadata sebagai data yang merinci karakteri data sumber, mendeskripsikan hubungannya serta menunjang penemuan dan penggunaannya yang efektif. Dengan meninjau berbagai definisi di atas maka metadata adalah data yang mendeskripsikan atribut sebuah sumber daya, mencirikan hubungannya, menunjang penemuannya dan penggunaannya secara efektif serta berada di lingkungan elektronik. Metadata biasanya terdiri atas himpunan unsur data, masing-masing elemen (unsur) memori (mendeskripsi) atribut sumber daya, manajemennya atau penggunaannya.

Pengertian metadata seperti diuraikan di atas mungkin masih kabur bagi banyak pustakawan. Namun kalau dikatakan bahwa beberapa skema metadata telah lama digunakan oleh pustakawan seperti metadata yang paling terkenal yaitu MARC dan INDOMARC maka

pustakawan mulai dapat membayangkan apa yang dimaksud dengan metadata. Contoh metadata lainnya ialah format katalog AACR2, daftar tajuk subjek (semacam Library of Congress Subject Headings) serta skema klasifikasi seperti DDC, UDC dan sejenisnya (Chowdhury 1999).

Setiap skema dirancang oleh pakar dalam bidang masing-masing dengan mempertimbangkan domain spesifik, kebutuhan sumber daya informasi dan persyaratan untuk pendeskripsian sebuah dokumen. Semua skema itu digunakan untuk akses bibliografi dan kontrol di perpustakaan, pusat dokumentasi, pusat analisis informasi dan lain-lainnya yang kesemuanya itu merupakan badan informasi (Sulistyo-Basuki 2000).

Secara umum ada 3 bagian yang digunakan untuk membuat metadata bagi sebuah paket informasi. Adapun ketiga bagian itu ialah (1) penyandian (encoding), (2) pembuatan deskripsi paket informasi bersama dengan informasi lain yang diperlukan untuk manajemen dan preservasi paket dan (3) penyediaan akses terhadap deskripsi tersebut. Karena keterbatasan waktu dan tempat maka makalah ini membahas tentang deskripsi saja, tidak mencakup penyandian maupun akses.

3. Fungsi metadata

Berdasarkan definisi metadata yang ada, maka dapat dibuat standar atau skema metadata berdasarkan provenans (asal usul), bentuk, fungsi, statistik penggunaan, syarat dan ketentuan penggunaan, data administratif, peringkat atau *rating* isi, kaitan atau hubungan data, data struktural dan sebagainya. Keputusan menyangkut bagian mana yang akan dicakup tergantung pada pemahaman disainer sistem mengenai fungsi primer skema metadata.

Salah satu fungsi utama metadata ialah penemuan sumber (*resource discovery*). Di lingkungan Ilmu Perpustakaan dan Informasi, kegiatan operasional ditekankan pada penelusuran, temu balik, penemuan (*discovery*) dan

akses ke sumber daya. Menurut Rao (1995) mengatakan bahwa kegunaan utama metadata ialah menunjang pemilihan, pemahaman, pen-dayagunaan dan pengingatan sumber dan isinya. Khususnya metadata memungkinkan mekanisme yang efektif untuk mengenali dan mengetahui lokasi data yang relevan dengan pemakai. Metadata memungkinkan pemakai untuk menentukan:

- (i) ketersediaan informasi (apakah objek informasi itu ada atau eksis? Di manakah letaknya? Berapakah yang tersedia? Apakah kesemuanya itu sama?)
- (ii) kegunaan informasi (apakah otentik? Apakah baik? Bagaimana pemakai dapat menentukan apakah berguna atau tidak?)

Bila komunitas Ilmu Perpustakaan dan Informasi memusatkan pada fungsi menemukan dan menelusur dan temu balik sumber daya, maka komunitas manajemen data berorientasi ilmu komputer memfokuskan pada aspek penggunaan data. Pengarsipan data dari segi ilmu komputer memerlukan skema untuk mendeskripsi struktur data logis atau konseptual dari semua objek atau maujud (entitas) yang berkaitan dengan arsip serta hubungan antara data (Strawman and Bretherton 1994). Strawman dan Bretherton berpendapat bahwa dengan konteks yang jelas dan terstruktur maka perbedaan antara data dengan metadata tidak ada lagi. Jadi menurut perspektif ini maka perbedaan antara metadata dengan data adalah penggunaannya saja. Fokus selanjutnya diarahkan ke tugas memberi batasan tentang konteks. Konteks ini meliputi berbagai persyaratan fungsional seperti fungsi administratif (misalnya autentifikasi pemakai dan mekanisme pembayaran), fungsi penentuan isi (misalnya analisis data untuk menunjang pemahaman tentang makna data), fungsi semantik sintaktik (misalnya pengembangan struktur rekod atau cantuman) dan reorganisasi data untuk penyajian dan visualisasi. Konteks yang berlainan menekankan fungsi yang

berbeda-beda. Hunter dan Springmeyer menyatakan bahwa fungsi dasar metadata ialah membantu manajemen data dan sistem penyimpanan dalam menyediakan akses yang lebih efisien ke himpunan data yang besar. Strebel (1994) mengatakan adanya 3 fungsi metadata yaitu (a) manajemen data; (b) akses data dan (c) analisis data.

Fungsi metadata juga dapat dikaitkan dengan aras sistem dan tingkat pemakai. Pada tingkat sistem, metadata dapat digunakan untuk memudahkan interoperasional dan keterbagian di antara berbagi sarana penemuan sumber daya. Berbagi data akan mempercepat penyelesaian proyek, meningkatkan pemanfaatan penelitian dan pengambilan keputusan serta mengurangi biaya dengan cara meminimumkan upaya duplikasi. Berpatungan data juga menunjang integrasi sumber daya Internet dan materi tercetak yang sudah diwakili dalam format terbaca mesin. Pada tingkat pemakai, metadata memudahkan kemampuan untuk menentukan (1) data apa yang tersedia; (2) apakah data tersebut memenuhi kebutuhan tertentu; (3) bagaimana memperolehnya dan (4) bagaimana mentransfernya ke sistem setempat.

3. *Cantuman¹ bibliografis, surogat, metadata*

Cantuman bibliografis (bibliographic record) merupakan istilah yang diterapkan terhadap paket informasi tak teruraikan seperti buku dan rekaman suara. Walaupun kini diterapkan juga terhadap cantuman untuk gambar hidup, rekaman suara, berkas komputer dsb., namun karena ada istilah bibliografi yang berasal dari kata *biblio* artinya buku, tetap saja istilah tersebut menimbulkan alergi bagi pihak nonpustakawan. Ada yang menggantikannya dengan cantuman surogat atau rekod surogat. Istilah tersebut dapat digunakan untuk cantuman atau rekod yang mewakili setiap jenis paket

informasi dalam setiap jenis sistem temu balik informasi. Pada makalah ini cantuman atau rekod surogat digunakan sebagai deskripsi dan akses terhadap isi sebuah cantuman metadata.

Cantuman surogat adalah penyajian karakteristik sebuah paket informasi. Karakteristik tersebut meliputi data deskriptif dan titik akses. Cantuman merupakan pemerian paket informasi dalam sistem temu balik informasi seperti katalog, indeks, bibliografi, mesin telusur (search engines). Paket informasi merupakan contoh informasi terekam seperti buku, artikel, kaset video, dokumen Internet atau himpunan "halaman elektronik", rekaman suara, jurnal elektronik dsb. Data deskriptif adalah data yang berasal dari paket informasi seperti judul, pengarang, edisi, tahun publikasi, catatan yang digunakan untuk mendeskripsinya. Titik akses adalah setiap istilah (kata, tajuk dsb) dalam cantuman surogat yang digunakan untuk menemubalik cantuman. Titik akses seringkali dikeluarkan dari data deskriptif, ditempatkan di bawah kontrol akses.

4. *Standar penyandian (encoding)*

Cantuman surogat dalam sarana bibliografi perlu disandi untuk manipulasi mesin bilamana cantuman surogat tersebut akan diletakkan di pangkalan data terpasang (online database). Cara untuk menyandikan itu antara lain ialah MARC (Machine Readable Catalogue) digunakan untuk menyandi katalog perpustakaan, SGML (Standard Generalized Markup Language) untuk menyandi teks, digunakan untuk menyandi cantuman, HTML (HyperText Markup Language) serta aplikasi SGML lainnya untuk keperluan WWW.

Cantuman surogat disandi (encode) dengan menambahkan tengara (tags), bilangan, huruf atau kata (yaitu kode) untuk mendiskrit butir informasi. Maka pada sandi MARC misalnya nama pengarang perorangan diberi tengara "100", sedangkan pada HTML didahului dengan "<author>" dan diikuti dengan "<author>".

¹ Cantuman merupakan terjemahan kata *records* yang banyak digunakan di kalangan pustakawan. Di lingkungan kearsipan istilah tersebut diterjemahkan menjadi rekod atau arsip dinamis.

4.1. Alasan penyandian (coding)

Penyandian cantuman surogat dilakukan karena berbagai alasan yaitu:

- (1) untuk keperluan paparan (display). Penyandian memungkinkan pemisahan masing-masing bagian cantuman (disebut ruas atau *field*), kemudian program komputer menulis masing-masing ruas sehingga setiap ruas berada dalam posisi tertentu di paparan sesuai dengan keinginan pembuat.
- (2) Untuk menyediakan akses bagi cantuman surogat. Penyandian memungkinkan pembuatan program penelusuran yang memungkinkan pemakai menelusur ruas tertentu. Misalnya bila ingin menelusur nama pengarang maka sandi MARC, ruas 100 dan 700 sudah cukup untuk mencari ruas pengarang perorangan, tidak perlu memeriksa ruas lainnya.
- (3) Memungkinkan integrasi berbagai bahasa dan aksara (*script*) yang dipaparkan dan ditelusur dalam berkas (file) yang sama. Badan informasi mengumpulkan informasi dalam banyak bahasa dan aksara. Bahasa yang tertulis dalam aksara nonLatin perlu diromanisasikan untuk memudahkan penjarannya.

4.2. Standar penyandian

Standar penyandian yang ada dewasa ini ialah MARC yang muncul dalam berbagai versi seperti USMARC, UNIMARC, INDOMARC, SGML (Standard Generalized Markup Language) mencakup DTD (Document Type Definition) terbagi lagi menjadi TEI (Text Encoding Initiative), HTML (Hype Text Markup Language), EAD (Encoded Archival Description) DTD, MARC DTD dan XML (Extensible Markup Language).

4.2.1. MARC

Machine Readable Cataloguing (MARC) diciptakan tahun 1968, kemudian namanya berubah menjadi USMARC setelah muncul MARC versi lain seperti CAN/MARC, UKMARC, DenMARC dsb. UNIMARC

(UNiversal MARC) dikembangkan tahun 1977 sebagai wahana pertukaran cantuman MARC antara berbagai badan bibliografi nasional. Untuk Indonesia dikembangkan INDOMARC oleh Perpustakaan Nasional RI.

4.2.2. DTD (Document Type Definition)

DTD merupakan sebuah aplikasi SGML. DTD dibuat dengan mengikuti peraturan *markup* yang terdapat pada standar SGML. DTD mendefinisikan struktur *tipe* dokumen tertentu

DTD memberi batasan atau definisi tentang:

- semua elemen mungkin merupakan bagian dari tipe dokumen tertentu.
- elemen nama dan apakah elemen tersebut terulangkan.
- bagaimana susunan penempatan elemen tersebut.
- isi elemen secara umum, bukan secara khusus.
- jenis *markup* yang dapat dihilangkan.
- atribut tengara (*tag*) serta nilai *default*nya.
- name entitas atau maujud (*entity*) yang dimungkinkan.

DTD dapat dibuat untuk sebuah dokumen. Dalam hal demikian DTD ditempatkan pada awal teks. Namun demikian pembuatan DTD memerlukan banyak waktu, karena itu lebih bermanfaat menciptakan DTD yang dapat digunakan untuk berbagai dokumen. DTD semacam itu terpisah dari teks yang merujuknya. Ada pula cara lain untuk membuat DTD yang dibuat khusus untuk aplikasi yang luas. Hal ini dilakukan dengan menempatkan pernyataan pada awal dokumen yang memberitahukan tengapa apa yang digunakan. Dokumen yang sesuai dengan SGML harus menyatakan DTD ekstern yang digunakan atau menggunakan DTD intern untuk dokumen yang bersangkutan.

DTD yang dibuat antara lain ialah :

- (1) TEI-DTD yaitu DTD untuk menyandi teks literer
- (2) HTML-DTD yaitu DTD untuk menyandi halaman *web*

- (3) EAD-DTD yaitu DTD untuk menyandi alat bantu temu kearsipan
- (4) MARC-DTD yaitu DTD untuk menyandi cantuman USMARC

4.2.3. TEI (*Text Encoding Initiative*)

TEI DTD dibuat untuk mengatasi kesulitan skema penyandian jamak yang digunakan untuk menyandi teks literer kuno dan atau kependitan (*scholarly*). Setelah disandi, dokumen tersebut tidak mudah saling ditukarkan. TEI memungkinkan membuat ciri khas sebuah teks secara eksplisit dalam sebuah format dan format tersebut memungkinkan pengolahan teks oleh berbagai program berlainan pada mesin yang berlainan pula. Teks dapat disajikan persis sebagaimana dalam bentuk aslinya.

TEI dapat digunakan untuk dokumen yang baru dibuat, terutama dalam hal penulis yang memiliki visi bagaimana teksnya terwujud. TEI semula digunakan untuk bidang Ilmu Pengetahuan Budaya namun kini meluas kebidang lain.

4.2.4. HTML (*HyperText Markup Language*)

HTML adalah himpunan kode yang dapat disisplam kedalam berkas teks untuk menunjukkan *typefaces* khusus, menyisipkan citra dan mengaitkan ke dokumen hypertext lainnya. HTML DTD dikembangkan untuk memungkinkan penciptaan halaman *web*. HTML merupakan bahasa *markup* dasar yang memungkinkan siapa saja menjadi penulis *web*. HTML memiliki kemudahan berupa pembuatan strukturyang sederhana, memaparkan citra serta kaitan antara dokumen. Pemakai dokumen bersandi HTML dapat navigasi melalui teks itu sendiri bilamana telah dibuat hubungan intern atau dapat bergerak dari satu teks ke teks lain dengan kaitan eksternal.

4.2.5. EAD (*Encoded Archival Description*) DTD.

EAD DTD digunakan sebagai alat bantu materi perpustakaan dan kearsipan, khususnya untuk membuat register dan inventaris. EAD DTD tidak merinci isi intelektual namun memberikan definisi arahan penyandian (*decoding designation*). EAD membantu kemampuan pertukaran alat bantu temu di kalangan depo arsip serta memungkinkan pemakai mengetahui koleksi di tempat lain.

4.2.6. MARC DTD

MARC DTD yang dihasilkan selama ini berjumlah lebih dari satu karena adanya format USMARC yang berlainan. MARC DTD memperlakukan cantuman MARC sebagai tipe dokumen tertentu. Memberi batasan elemen yang dapat muncul pada sebuah cantuman MARC, mengatur bagaimana penengaraannya dan diwakili dalam penyandian SGML. MAEC DTD berujuan agar semua cantuman MARC dapat secara otomatis diterjemahkan ke SGML dan cantuman MARC tersandikan dalam SGML dapat diterjemahkan kembali dengan mudah menjadi sebuah cantuman MARC. Hal tersebut dimungkinkan karena setiap ruas pada cantuman tersandi dalam SGML berisi semua tengara, indikator dan kode subruas MARC di samping isi ruas.

4.2.7. XML (*Extensible Markup Language*)

XML merupakan sebuah subhimpunan dari SGML. HTML dikritik karena tidak mampu memberikan aplikasi yang diminta akibat sifatnya yang sangat sederhana. Ada yang berpendapat bantuan SGML di *web* akan dapat membantu. Namun SGML bersifat rumit, memiliki ciri yang memerlukan pemrograman yang rumit dan panjang. Upaya lain berupa pengembangan *plug-in* sehingga pemakai dapat memindah dan memasangnya.

Bahasa pemrograman Java dapat meningkatkan fungsionalitas HTML namun untuk keperluan itu pemakai harus mampu menciptakan halaman *web* di dalamnya. Maka dikembangkan XML untuk mengatasi masalah tersebut. XML tidak memuat beberapa ciri

SGML, memiliki metode untuk membaca teks nonASCII (seperti teks Cina), mengimplementasi UNICODE yang di dalamnya termasuk ASCII/

Salah satu aplikasi XML ialah RDF (Resource Description Framework) untuk menciptakan metadata. RDF menyediakan lebih dari satu jenis data (misalnya data tipe katalog; peringkat isi untuk menunjukkan misalnya apakah isi sesuai dengan anak-anak atau tidak indikasi hak milik intelektual dsb). Kini berkembang upaya pengaitan Dublin Core dengan RDF.

5. Penciptaan isi cantuman surogat

Cantuman surogat dibuat dengan cara memiliki butir informasi yang penting dari sebuah paket informasi (misalnya pengarang, judul, tahun dsb), kemudian menentukan karakteristik tertentu tentang paket informasi (misalnya besaran, syarat keterperolehan atau terms of availability). Butir informasi tersebut kemudian disusun menurut urutan tertentu berdasarkan ketentuan peraturan atau konvensi untuk deskripsi. Peraturan atau kebiasaan ini dibuat oleh bermacam-macam komunitas guna memenuhi kebutuhan komunitas tersebut akan deskripsi paket informasi. Maka akan terdapat bermacam-macam alat pembuatan cantuman surogat seperti :

- (1) *ISBD (International Standard Bibliographic Description)*
- (2) *AACR2 (Anglo-American Cataloguing Rules, Second Edition, 1988 revision)*
- (3) *APPM (Archives, Personal Papers, and Manuscripts, 1989)*
- (4) Dublin Core
- (5) TEI (Text Encoding Initiative) Headers
- (6) GILS (Government Information Locator Service)
- (7) FGDC (Federal Geographic Data Committee) Content Standard for Digital Geospatial Metadata (CSDGM)
- (8) VRA (Visual Resources Association) Core Categories for Visual Resources

(9) EAD (Encoded Archival Description)

4.1.-4.3. ISBD, AACR2, APPM.

Tentang ISBD, AACR2 maupun APPM tidak akan dibahas di sini karena diasumsikan semua sudah tahu dan atau dianggap terlalu teknis ke perpustakaan dan kearsipan. Perpustakaan memiliki tradisi yang panjang dalam pengembangan sistem informasi. Tradisi ini terpusat pada pengorganisasian dan penyediaan akses ke maujud atau entitas yang memuat informasi seperti buku dan materi tercetak. Organisasi dan akses disediakan dengan sarana berupa deskripsi bibliografi, analisis subjek dan skema klasifikasi. Dengan menggunakan deskripsi bibliografi, analisis subjek dan skema klasifikasi maka perpustakaan dapat mendeskripsi, mengidentifikasi, menentukan lokasi dan menemu balik semua maujud memuat informasi. Maka muncullah Anglo American Cataloguing Rules, International Standard Bibliographic Description untuk deskripsi bibliografis; *Library of Congress Subject Headings*, *Library of Congress Subject Cataloging* untuk analisis subjek; Dewey Decimal Classification dan Universal Decimal Classification untuk klasifikasi. Berbekal sarana tersebut perpustakaan mampu mengimplementasi sistem pengaturan informasi berbasis surogat. Surogat ini berbentuk katalog, tajuk subjek dan notasi klasifikasi. Prinsip yang digunakan untuk sistem tersebut menekankan pada deskripsi, lokasi dan temu balik kontener informasi. Misalnya tiga objek katalog sebagaimana dikemukakan oleh Cutter (1904, 12) tetap tidak berubah walaupun dalam lingkungan digital. Adapun ketiga tujuan itu ialah:

1. Memungkinkan seseorang menemukan sebuah buku;
2. Menunjukkan apa yang dimiliki perpustakaan dan
3. Membantu dalam pemilihan sebuah karya.

Unit dasar sistem pengkatalogan perpustakaan adalah surogat yang mewakili objek informasi.

Manfaat sebuah surogat ditentukan oleh seberapa jauh:

- a. dapat diketahui lokasinya berdasarkan pengarang, judul, subjek
- b. identifikasi dan mendeskripsikan objek secara tepat; dan
- c. identifikasi dan mendeskripsikan lokasi mutakhir objek tersebut.

4.4. Dublin Core

Dublin Core merupakan singkatan dari Dublin Metadata Core Element Set diciptakan untuk membuat himpunan elemen yang disepakati secara internasional yang dapat "diisi" oleh pembuat dokumen elektronik. Peserta pertemuan Dublin Core berasal dari berbagai bidang (seperti penerbit, spesialis komputer, pustakawan, produsen perangkat lunak) dari berbagai negara termasuk Indonesia serta dibentuk Dublin Core dalam beberapa bahasa termasuk bahasa Indonesia (Prabawa, 1998).

Agar sebuah dokumen memiliki kemampuan swaindeks (self-indexed) maka sekumpulan elemen metadata perlu diidentifikasi sehingga setiap pembuat atau pencipta dokumen elektronik dapat menerapkannya pada dokumen elektronik yang diciptakannya. Dengan berpegang pada sasaran sederhana itu maka muncul upaya menyusun himpunan deskripsi data untuk dokumen elektronik.

Salah satu himpunan deskripsi elemen metadata itu adalah *Dublin core* (http://purl.org/metadata/dublin_core) yang dikembangkan oleh OCLC yang berpusat di Dublin, Ohio. Dublin core menentukan 15 elemen metadata berupa:

1. judul
2. pencipta atau pembuat (*creator*)
3. subjek (katakunci, kosakata teekendali dan klasifikasi)
4. deskripsi (abstrak dan deskripsi isi)
5. penerbit
6. penyumbang (terkecuali pencipta)
7. tahun
8. tipe (kategori sumber)
9. format (HTML, Postscript)

10. pengenal atau *identifier* (URL, untaian atau nomor yang digunakan untuk mengenali sumber)
11. sumber (darimana asal usul sumber daya)
12. bahasa
13. hubungan atau kaitannya dengan sumber daya lain
14. cakupan (spasial dan atau karakteristik sumber daya)
15. hak (hubungannya dengan pemberitahuan hak cipta)

Elemen data Dublin core hanya bersifat deskriptif yang memiliki nilai intrinsik. Dengan demikian Dublin core mampu mengurangi atau menghilangkan penggunaan rujukan eksternal (seperti pada peraturan pengkatalogan atau *authority files*). Elemen data tersebut dapat perluasan sehingga meliputi informasi khusus tambahan; bersifat bebas dari sintaksis, terulang-kan dan dapat diubah melalui *qualifiers* (keterangan tambahan) untuk memperluas maknanya sampai melewati makna di luar makna yang lazim.

Untuk memungkinkan pembuat cantuman menerapkan metadata, maka dibuatkan mekanisme untuk menyatukan data dengan dokumen HTML (Hypertext mark up language).

4.5. Text Encoding Initiative (TEI) Headers

TEI adalah sebuah SGML DTD yang dibuat untuk menyediakan penyandian bagi teks kuno, literer atau kependitan (scholarly) sehingga versi yang tersandi dapat saling tukar dengan mudah. TEI Header dibuat sehingga terdapat metadata sebagai bagian dari berkas teks.

Pembuatan TEI Header bertujuan menyediakan sumber informasi untuk pengkatalogan. Standar REI Header banyak mengambil dari AACR2 sehingga isi TEI Header mirip dengan AACR2.

TEI Header memiliki 4 bagian yaitu deskripsi berkas, deskripsi penyandian, deskripsi profil dan deskripsi revisi. Deskripsi berkas diwajibkan, berisi deskripsi bibliografi dari teks. Di dalamnya termasuk judul; pengarang; informasi publikasi dan deskripsi sumber. Deskripsi

sumber merupakan deskripsi sumber asli dan dari sumber asli dibuatkan teks elektronik. Deskripsi penyandian menjelaskan peraturan atau keputusan editorial yang digunakan dalam transkripsi teks, misalnya variasi ejaan yang digunakan. Deskripsi profil berisi titik akses, misalnya nama lain selain nama pengarang, informasi bahasa, titik akses subjek dan notasi klasifikasi. Deskripsi revisi berisi catatan tentang setiap perubahan yang telah dilakukan terhadap teks termasuk bilamana perubahan dilakukan serta oleh siapa.

4.6. *GILS (Government Information Locator Service) Records*

GILS muncul pada tahun 1994 setelah U.S. Congress menelurkan *paperworks Reduction Act* dan *Office of Management and Budget* mewajibkan kantor federal menyediakan metadata umum tentang pemilikan informasi mereka. GILS mensyaratkan semua badan Federal menyediakan cantuman yang disebut "locator" terhadap setiap sumber daya informasi yang dimilikinya. Banyak kantor Federal hanya membuat sebuah cantuman singkat yang mendskripsikan semua informasi sebagai keseluruhan.

Format metadata GILS bersifat rumit sehingga memungkinkan mencakup informasi singkat sampai terinci. Unsur GILS disebut "data element" meliputi butir seperti Judul, pengarang, subjek, tahun publikasi, tanggal terakhir dimodifikasi, periode waktu dari isi, domain spasial, pembatasan penggunaan, point-of-contact dan abstrak.

Sebuah cantuman GILS memberikan deskripsi yang memberikan informasi paling sedikit mencakup:

- informasi apa yang tersedia dan mengapa informasi itu dibuat
- bagaimana cara penyediaan informasi
- siapa yang perlu dikontak untuk memperoleh informasi lebih lanjut
- kemungkinan kaitan (link) langsung ke informasi itu sendiri

4.7. *FGDC (Federal Geographic Data Committee) Content Standard for Digital Geospatial Metadata (CSDGM)*

Diwajibkan bagi badan Federal AS sejak tahun 1994 yang mewajibkan mereka merekam semua data geospasial baru yang mereka kumpulkan atau hasilkan. Standar tersebut bertujuan menghasilkan terminologi dan definisi umum untuk menatadat tentang data geospasial digital.

Standar FGDC merupakan standar isi, tidak mengatur tentang tataletak maupun skema penyandian. Standar tersebut menyediakan jalan bagi pemakai untuk mengetahui sumber daya apa saja yang tersedia, apakah memenuhi kebutuhannya, di mana menemukan data dan bagaimana mengakses data.

Standar FGDC terbagi atas informasi identifikasi; informasi mutu data; informasi pengaturan dataspasial; informasi rujukan spasial; informasi maujud (entity) dan atribut; informasi distribusi; informasi rujukan metadata; informasi sitiran; informasi waktu dan informasi kontak.

4.7. *VRA (Visual Resources Association) Core Categories for Visual Resources*

Kategori inti (core categories) dimaksudkan sebagai panduan untuk mendskripsi dokumen visual yang melukiskankarya seni, arsitektur dan artefak atau struktur dari material, populer dan budaya rakyat (folk culture). Kategori deskripsi kerja terbagi atas 23 kategori sedangkan kategori deskripsi dokumen visual terdiri dari sembilan buah.

Work Description Categories

Work type	Repository name	Movement
Title	Repository Place	Nationality
Measurements	Repository Number	Culture
Material	Current Site	Subject
Technique	Original Site	Related Work
Creator	Style	Relationship type
Role	Period	Notes
Date	Group	

Visual Document Description Categories

Visual Document Type	Visual Document Owner
----------------------	-----------------------

Visual Document Format	Number Visual Document View Description
Visual Document Measurements	Visual Document Subject
Visual Document Data	Visual Document Type
Visual Document Owner	

4.8. EAD (*Encoded Archival Description*)

EAD dibuat khusus untuk menyandi alat bantu temu terutama untuk materi kearsipan. Karena panduan alat bantu temu sudah ada, maka EAD tidak berisi ketentuan tentang isi. EAD memiliki tajuk yang sangat tergantung pada TEI Header. Tajuk tersebut merupakan metadata tentang metadata pada cantuman EAD. Isi yang ditempatkan pada EAD Header berdasarkan peraturan pada *APPM (Archives, Personal Papers, and Manuscripts)*. Kini mulai banyak koleksi arsip yang dikatalog dengan menggunakan APPM/MARC.

4.9. *Crosswalks*

Crosswalks adalah instrumen visual untuk menunjukkan bagian sebuah standar metadata yang cocok dengan standar lain. Contohnya ialah "Dublin Core/MARC/GLIS Crosswalk" yang memetakan Dublin Core pada USMARC dan GILS serta "'onticello Electronic Library: Dublin Core Element Set Crosswalk'" yang memetakan Dublin Core pada EAD, GILS dan USMARC.

4.10 *On-the-fly records*.

On-the-fly records adalah cantuman yang diciptakan secara elektronik untuk keperluan segera. Cantuman tersebut akan berubah dalam waktu singkat atau dalam hitungan hari atau minggu. Pada beberapa katalog yang telah memiliki "web interfaces", cantuman MARC akan dikonversi menjadi cantuman tersandi HTML *On-the-fly* atau saat itu juga. Hal itu terjadi manakala pemakai dalam penelusuran menyebutkan pengarang, judul dan atau tahun. Bila pemakai mengklik nomor 4 pada daftar,

maka cantuman MARC dimunculkan dalam layar melalui konverter HTML

5. *Internet dan 2 jenis ancangan (approach)*

Mulai awal 1990an Internet digunakan oleh masyarakat termasuk perpustakaan. Internetpun mulai digunakan di Indonesia pada awal tahun. Perpustakaan Indonesia pun mulai memasang dan menggunakan internet. Karena salah satu fungsi perpustakaan menyediakan informasi bagi pemakainya terutama dari koleksinya maka timbul pertanyaan bagaimana menyediakan informasi tentang koleksi sebuah perpustakaan atau katalog perpustakaan di Internet.

Pemikiran itu timbul karena data dari katalog harus ditaruh di situs Web 1998) sedangkan di segi lain dokumen elektronik perlu swaindeks (*self-indexed*) sebagai lawan pengkatalogan penambahan, tengara dan nilai pengindeksan yang dilakukan oleh lembaga pengkatalogan dan pengindeksan termasuk perpustakaan. Dengan kata lain perlu sebuah format katalog yang akan ditempatkan di situs web.

Kegiatan intelektual yang ada di perpustakaan, badan informasi maupun di internet merupakan kegiatan intelektual. Salah satu tugas utama kegiatan intelektual adalah kemampuan untuk melokalisir, mengenali, menemukbalik dan memanipulasi informasi. Tugas ini dilakukand dengan berbagai cara serta didekati oleh berbagai jenis komunitas intelektual. Dalam lingkungan digital, ada dua ancangan utama yaitu ancangan pengawasan bibliografis berorientasi ilmu perpustakaan dan ancangan manajemen data berorientasi ilmu komputer.

Walaupun kedua ancangan tersebut memiliki teori dan praktek yang berlainan, kedua pendekatan menggunakan skema metadata untuk memudahka penyelesaian tugas aktivitas intelektual. Filosofis kedua ancangan memiliki penekanan dan ciri yang berbeda-beda yang mencerminkan sejarah dan konteks kedua pendekatan. Di dalam era digitalisasi objek informasi yang digunakan untuk mengolah, menyimpan, temubalik, penyebaran dan penggunaan objek informasi, maka kedua ancangan tersebut

memodifikasi skema mereka untuk menyesuaikan diri dengan lingkungan digital. Dalam proses penyesuaian dan penjelajahan ternyata antara kedua bidang bergerak saling mendekati.

5.1. *Ancangan pengawasan bibliografis berorientasi ilmu perpustakaan.*

Dalam bahasa Inggris dikenal sebagai *library science oriented bibliographic control approach*. Sebagaimana telah diuraikan di atas, aktivitas intelektual di dunia perpustakaan sudah memiliki tradisi yang lama, mulai dari berbagai pedoman deskripsi bibliografi, skema klasifikasi, daftar tajuk subjek yang digunakan dalam sistem informasi berbasis surogat. Surogat adalah perwakilan objek informasi yang diwujudkan dalam deskripsi bibliografi, notasi klasifikasi, tajuk subjek. Sebagai contoh buku *Pengantar Ilmu Perpustakaan dan Informasi* diwujudkan dalam bentuk katalog, notasi DDC 020 atau tajuk subjek *LIBRARY AND INFORMATION SCIENCE*. Kesemuanya itu merupakan surogat dari objek informasi buku *Pengantar Ilmu Perpustakaan dan Informasi*. Hal yang sama berlaku pula untuk objek informasi lainnya.

Kesadaran akan perlunya berbagai informasi secara nasional dan internasional di antara lembaga informasi telah menghasilkan berbagai standar, peraturan (*rules*) dan ketentuan seperti ISBD, MARC untuk meningkatkan interoperasionalitas dan penggunaan sistem informasi. Di Indonesia hal ini diwujudkan lebih lanjut misalnya terjemahan ISBD, INDO-MARC, tambahan skema klasifikasi untuk Islam, sejarah, geografi dll.

5.2. *Ancangan manajemen data berorientasi ilmu komputer*

Dalam bahasa Inggris disebut *computer science oriented data management approach*. Komunitas komputer juga memiliki tradisi yang lama menyangkut manajemen dan organisasi data. Hal ini diwujudkan dalam berbagai sistem simpan dan temu balik informasi terkomputer baik untuk data tekstual maupun data relasional.

Tujuan sistem tersebut untuk menyimpan, akses dan mendayagunakan data secara efektif, menyediakan pengamanan data, berbagi data dan berbagai fungsi keterpaduan data. Untuk itu dikembangkanlah berbagai jenis sistem simpan dan temubalik data untuk berbagai keperluan yang digunakan oleh komunitas yang berlainan seperti lembaga niaga, ilmu pengetahuan, pendidikan, kesenian dll. Bila arsip data tersebut menjadi sangat besar, terdistribusi dan terdiversifikasi, maka data yang dalam jumlah besar masalah representasi dan pemetaan data yang diatasi dengan mekanisme struktur data dan interogasi data yang kompleks. Berbagai model dan arsitektur data diajukan untuk memecahkan masalah tersebut di berbagai bidang.

5.3. *Data elektronik di ruang cyber (cyberspace)*

Internet yang semula digunakan untuk kepentingan militer, kemudian pada akhir tahun 1980an digunakan untuk kepentingan sipil termasuk untuk kepentingan perpustakaan. Internet merupakan saluran penyebaran informasi yang menggabungkan kombinasi yang secara tradisional dilakukan oleh perpustakaan dan arsip data yang dilaksanakan oleh manajemen data berbasis komputer. Namun ada perbedaan antara internet dengan komputer dan perpustakaan; perbedaan tersebut terletak pada arsitektur data. Perpustakaan dan arsip data pada komputer secara primer merupakan sistem simpan dan temu balik data sedangkan secara sekunder kedua-duanya merupakan media komunikasi. Sebaliknya, Internet secara primer merupakan media komunikasi dan secara sekunder merupakan sistem simpan dan temu balik informasi. Namun demikian dengan perkembangan teknologi digital, batasan antara fungsi primer dan sekunder antara perpustakaan dan arsip data di satu pihak dengan Internet di segi lain semakin kabur. Perpustakaan dan arsip data di komputer semakin mengandalkan pada penggunaan Internet sebagai sistem penghantaran informasi.

Di samping perbedaan di atas masih ada lagi perbedaan antar perpustakaan dan arsip data

dengan Internet. Perpustakaan tradisional dan arsip data tidak menggunakan model klien/server sedangkan Internet menggunakan model klien/server sehingga pada Internet lebih banyak memberikan kontrol pada pemakai atau pemakai akhir (end-user). Server memberikan beberapa wewenang kontrol pada pemakai yang dilakukan oleh pemakai tanpa harus melepaskan tanggung jawab server. Ini dilakukan dengan kerjasama antara disainer sistem, pemakai sistem untuk menghasilkan standar antarmuka yang menghasilkan kontrol lebih besar oleh pemakai tanpa perlu mengorbankan efisiensi penelusuran dan temu balik data. Kelompok standar ini dikenal sebagai standar interoperasional jaringan.

Perbedaan lain antara perpustakaan dan arsip data dengan Internet terletak pada penyimpanan data. Di perpustakaan dan arsip data, data disimpan secara sentral sedangkan pada Internet penyimpanan tidak terpusat melainkan terdistribusi atau teragih melalui interoperasional jaringan komputer. Karena itu sumber daya informasi memiliki sifat :

- (1) tidak tetap dan stabil (berbeda dengan dokumen cetak dan materi fisik lainnya di perpustakaan) serta tidak aman (berbeda dengan data pada arsip data);
- (2) tidak diseleksi dan dikumpulkan (berbeda dengan materi yang dipilih dan diperoleh sebagai koleksi perpustakaan atau dikumpulkan dan diperoleh di arsip data);
- (3) tidak terorganisir secara sentral untuk keperluan akses (misalnya dikatalog untuk koleksi perpustakaan) atau diindeks (sebagai tabel atribut data atau indeks data di arsip data).

Disainer sistem mencoba mengatasi masalah temu balik sumber Internet dengan mengembangkan mesin telusur (*search engines*). Mesin telusur ini bermanfaat bila pemakai memahami dasar mekanisme kerjanya namun dengan semakin meningkatnya jumlah sumber elektronik terpasang (*online electronic resources*) maka efektivitas dan efisiensi mesin telusur akan merosot tanpa memahami meta kontrolnya.

Mesin telusur Internet memiliki keterbatasan dalam menghadapi masalah seperti:

- (a) bagian manakah dari objek informasi yang perlu dipayar atau dipindai (*scanned*), apakah indeks, teks lengkap, tajuk dsb;
- (b) bagian manakah dari luaran penelusuran (*search output*) yang perlu dipaparkan, misalnya indeks, judul pertama, 100 kata pertama, atjuk, ringkasan dsb;
- (c) konteks atau kriteria manakah yang perlu dipertimbangkan untuk pembobotan (*weighting*) dan perluasan istilah tanya (*queiry terms*).

6. Ancangan terhadap format metadata

Standar untuk interoperasional jaringan merupakan bidang disainer sistem yang kurang memerlukan masukan dari profesional informasi semacam pustakawan, arsiparis, dokumentalis dll. Standar tersebut mengupayakan agar arsitektur klien/server beroperasi secara efektif dan efisien. Ini diartikan bahwa komunikasi berlangsung tanpa gangguan.

Di segi lain pengembangan standar temu balik informasi pada aras dokumen memerlukan kerjasama berbagai bidang seperti disainer, sistem, penyedia data, spesialis informasi bibliografis (pustakawan) dan spesialis pengkodean teks elektronik. Dengan kerjasama itu maka identifikasi, lokasi, temubalik, manipulasi dan penggunaan informasi dijital dan jasa elektronik yang disimpan di ruangsiber (*cyberspace*) atau terletak di ruang siber atau hubungan melalui ruang siber dipermudah. Untuk keperluan itu digunakan standar yang dikenal sebagai standar metadata.

Smith (1996) menjelaskan katakarakteristik metadata yang dioperasionalkan dalam konteks perpustakaan tradisional sebagai berikut:

- (1) menyediakan karakterisasi objek informasi individual dalam koleksi perpustakaan. Objek individual informasi dapat berupa

buku, majalah, disertasi, laporan penelitian dan sejenisnya.

- (2) Disimpan sebagai isi katalog perpustakaan. Ini mencakup pengarang, judul, deskripsi fisik, subjek, notasi klasifikasi, tajuk subjek dan sebagainya.
- (3) Digunakan terutama untuk membantu pemakai dalam mengakses objek informasi yang diminati pemakai. Akses ini mungkin berupa pengarang atau judul atau subjek atau kombinasi beberapa aspek, terpulang pada pemakai.

Kesemuanya itu pada perpustakaan tradisional disimpan di katalog perpustakaan atau bagi yang sudah menggunakan komputer disimpan sebagai pangkalan data (*database*). Dengan berkembangnya teknologi digital maka karakteristik di atas diubah dengan cara berikut :

- (1) organisasi meta-informasi dalam bentuk fisik dapat diganti dengan struktur organisasi yang lebih luwes namun dalam bentuk elektronik;
- (2) organisasi fisik tunggal dari sebuah koleksi objek informasi (seperti buku, majalah, disertasi dsb) dapat digantikan dengan organisasi objek informasi jamak (*multiple*) dan logis;
- (3) adanya objek informasi dalam bentuk digital memungkinkan penggunaan teknologi digital untuk keperluan identifikasi dan ekstraksi informasi.

Pendapat Smith mencerminkan kajian utama metadata yaitu pemodelan data. Maka muncul kedua pendekatan (dari ilmu komputer dan ilmu perpustakaan) untuk mengembangkan standar metadata untuk pemodelan data (atau objek informasi). Kedua ancangan itu dikembangkan melalui :

- (1) disainer sistem arsip data yang beranjak dari ilmu komputer bekerja dari basis pengetahuan yang berkembang dalam tradisi manajemen data; dan
- (2) perpustakaan dan lembaga yang berkaitan dengan perpustakaan, beranjak dari ilmu perpustakaan, bekerja dari basis pengetahuan yang dikembangkan dalam tradisi klasifikasi, katalogisasi dan pengindeksan dokumen.

Pada kelompok 1, fokus disainer sistem tertuju pada penggunaan data seperti dikatakan Gritton (1994) metadata mewakili informasi yang menunjang penggunaan data sejak masa penciptaan melalui masa pendayagunaannya. Dalam ancangan ini, setiap informasi tambahan yang dianggap berguna seperti deskripsi isi, pembatasan akses data administratif yang dianggap dapat memperbaiki penggunaan data yang tersimpan dianggap sebagai metadata yang baik.

Pada kelompok kedua, mekanisme dan sarana pengawasan bibliografis yang berorientasi pada perpustakaan diterapkan untuk mengkatalog sumber daya Internet. Mekanisme dan sarana pengawasan bibliografi yang digunakan meliputi:

- (a) *authority control* untuk mempertahankan konsistensi dalam titik akses;
- (b) entri analitis untuk bagian dari objek informasi yang komprehensif;
- (c) INDOMARC dengan ruas tetap dan tidak tetap, subruas terulangkan dan tidak terulangkan untuk format cantuman.

INDOMARC sebagai MARC yang disesuaikan untuk kepentingan nasional sayangnya belum pernah membahas hal tersebut sehingga sebenarnya pada butir c di atas contoh yang lebih tepat adalah USMARC. Upaya yang dilakukan pada kelompok kedua antara lain ialah MIT Library (Xu 1996) dan OCLC Cataloging Internet Resources Project "interCat" (Jul 1995). Bila pembaca memperhatikan bahwa tujuan utama katalogisasi perpustakaan adalah lokasi dan kolokasi maka hal tersebut diikuti oleh kedua

proyek tersebut adalah fungsi penemuan sumber daya sebagaimana dikemukakan dalam ruang lingkup Dublin Core (Weibel et al 1995) ...only those elements necessary for the discovery of the resource were considered. It was believed that resource discovery is the most pressing need...

6. Metadata dan standar metadata

Secara umum definisi metadata adalah data tentang data. Dalam definisi tersebut pembaca menemukan kata *tentang* yang perlu kita pahami dalam konteks Ilmu Perpustakaan. Dari segi pengawasan bibliografi, fokus "tentang" adalah karakterisasi sumber data (Smith 1996; Weibel et al. 1995) untuk mengenali atau mengidentifikasi lokasi objek informasi dan memudahkan kolokasi isi subjek. Karena itu metadata adalah setiap informasi yang merekam karakterisasi dan hubungan sumber data atau himpunan elemen data yang dapat digunakan untuk mendeskripsikan dan mewakili objek informasi.

Di atas telah dibahas adanya dua ancangan yaitu ancangan berbasis ilmu perpustakaan dan ilmu komputer. Pada perspektif manajemen data berorientasi ilmu komputer, fokus "tentang" adalah memperkuat penggunaan data sumber (Gritton 1994; Strebel et al 1994; Strawman & Bretherton 1994). Metadata dari perspektif ilmu komputer adalah setiap informasi yang menunjang penggunaan efektif akan data, termasuk informasi yang dapat memudahkan manajemen data (misalnya otentifikasi data, urutan data, jenis data, indikator ruas utama), akses data (misalnya cakupan, parameter laporan) serta analisis data (misalnya format untuk *data mining*², visualisasi) (Rao et al 1995).

Walaupun fokus kedua ancangan itu berbeda, al tersebut tidaklah berarti bahwa kedua ancangan itu tidak tersetarakan (*incompatible*) maupun eksklusif mutual. Dalam lingkungan Internet, komunitas pemakai demikian ber-

variasinya sehingga hampir tidak mungkin mengenali penggunaan predominan tunggal di antara berbagai penggunaan Internet. Karena itu skema metadata bersifat luwes sehingga mampu memuaskan pemakai sebanyak mungkin.

Karena metadata telah berkembang dalam berbagai konteks serta berbagai keperluan maka terdapat berbagai format seperti Internet Anonymous FTP Archive (IAFA), Text Coding Initiative (TEI).

6.1. Format metadata untuk sumber daya Internet

Hingga saat ini belum ada standar internasional untuk metadata. Dempsey & Heery menyebutkan adanya beberapa skema metadata untuk objek informasi digital, mulai dari Dublin Core hingga ke Text Encoding Initiative (TEI). Di bawah ini perbandingan antara Dublin Core, IAFA *templates*, WWW Semantic Header, URC (Uniform Resources Characteristics atau Uniform Resources Citation), OCLC Intercat project yang menggunakan format USMARC dan dengan demikian secara tidak langsung dapat dibandingkan dengan INDOMARC, TEI (Text Encoding and Interchange).

² *Data mining* artinya akses data dari basis data yang sangat besar

Tabel 1 Perbandingan atribut format Metadata

	Dublin Core	URC	Semantic header	US MARC	IAFA Templates	TEI header
INTRINSIK						
Subjek	+		+	+	+	
Judul	+	+	+	+	+	+
Pengarang	+	+	+	+	+	+
Penerbit	+	+	*	+	+	+
Tempat publikasi		+	*	+		+
Agen lain	+		+	+		*
Tahun	+	+	*	+		
Tipe objek	+		*	+		
Bentuk	+	+		+		
Pengenal (URN, ISBN)	+	+	+	+	+	+
Hubungan	+	+	*	+		*
Sumber	+			+	+	+
Bahasa	+		*	+	+	*
Cakupan	+		*	+		
Abstrak		+	*			
Versi (edisi)		+	*			*
Catatan (anotasi)		+	+			
Tandatangan		+	+			
Klasifikasi				+		*
Klasifikasi (tingkat pengamanan)			*			
Kata kunci			+		+	*
EKSTRINSIK						
Persyaratan sistem			*		+	
Modus akses				+	+	
Ketersediaan					+	*
Biaya			*		+	*
Kontrol			*		+	
Keluasan (besaran)			*		+	*
Deskripsi penyandian (encoding)					+	*
Deskripsi revisi			*		+	*

+ wajib *opsional

Sumber: Ng, Park, Burnett (1997) dengan ubahan oleh penulis

Mengklasifikasi format metadata berdasarkan asal usul (misalnya apakah skema di atas berasal dari ilmu perpustakaan ataukah ilmu komputer) dapat menimbulkan kontroversi karena adanya kasus taksa (ambiguitas) dan upaya menggabungkan dalam bentuk kolaborasi antara kedua ilmu. Misalnya ada yang menganggap bahwa Dublin Core berasal dari ilmu perpustakaan karena keikutsertaan salah satu konsorsium terbesar di dunia perpustakaan yaitu OCLC dalam penyusunannya. Seusai lokakarya pertama tentang Dublin Core disusul

dengan pertemuan gabungan bersama American Library Association sehingga kental sekali warna kepeustakawanannya. Akan tetapi lokakarya Dublin Core juga diselenggarakan oleh National Centre for Computer Applications. Masalah Dublin Core yang berkaitan dengan INDOMARC justru dilakukan oleh seorang dosen Indonesia berlatar belakang ilmu komputer (Prabawa 1998).

TEI secara sekilas menampakkan warna ilmu komputer karena struktur definisi jenis dokumen serta pernyataan sistemnya, namun isi

tajuk bebas TEI sangat dipengaruhi oleh ilmu perpustakaan.

INDOMARC yang penyusunannya banyak bersumber pada USMARC sedikit banyak dapat dibandingkan dengan format metadata lainnya di Tabel 1. Pada Tabel 1 banyak elemen yang tidak terdapat pada USMARC. USMARC menyediakan lebih banyak ruas untuk judul berbagai bentuk dalam ruas 130, 210, 220, 240, 245, 246, 247, 630, 730 dan 740) namun kurang mencakup fungsi data. Dibandingkan dengan USMARC maka ruas pada INDOMARC lebih sedikit. Dengan kata lain USMARC dapat disebut kaya dalam penyediaan ruas format namun kurang menunjang fungsi akses data dan penggunaan data.

Elemen metadata pada USMARC secara garis besar terbagi atas 2 jenis yaitu intrinsik dan ekstrinsik. Intrinsik artinya elemen metadata yang berkaitan dengan identifikasi dan penemuan sumber daya sedangkan ekstrinsik berhubungan data administrasi dan non-bibliografis lainnya.

Bagi Indonesia sudah dirasakan perlunya pengembangan INDOMARC sehingga mencakup lebih banyak materi perpustakaan. Saat ini INDOMARC hanya digunakan untuk monograf, terbitan berseri sementara dalam AACR 2 masih terdapat bentuk materi perpustakaan yang lain seperti materi nonbuku, manuskrip, berkas komputer, materi kartografis, musik tercetak dll. Pengembangan lebih lanjut memerlukan kerjasama antara Perpustakaan Nasional, pustakawan praktisi di lapangan serta dari kalangan akademis.

Juga mulai dirasakan perlunya pembentukan forum tentang format metadata misalnya yang berkaitan dengan Dublin Core. Di sini penulis meyarankan topik Dublin Core karena Dublin Core merupakan format yang relatif sederhana, sudah ada forum yang membahasnya serta adanya tesis tentang Dublin Core dalam kaitannya dengan kebutuhan Indonesia yang justru dilakukan oleh pengajar dari ilmu komputer dari Indonesia. Di negara Asia Tenggara seperti Thailand dan Singapura sudah terbentuk

forum sejenis sehingga diharapkan sumbangan dari pustakawan Indonesia.

Di segi pendidikan pustakawan, pada mata kuliah Organisasi Informasi atau Katalogisasi dan Klasifikasi perlu dimasukkan komponen metadata agar mahasiswa siap melaksanakan tugas organisasi informasi di Internet serta sebagai pemutakhiran bahan kuliah. Bila perlu nama mata kuliah diganti menjadi Arsitektur Informasi!

7. Penutup

Objek informasi dideskripsi, disimpan, ditemubalik dan didayagunakan oleh perpustakaan dan badan informasi lainnya. Kegiatan intelektual kepustakawanan tersebut dituangkan dalam bentuk deskripsi fisik, klasifikasi, tajuk subjek. Dengan berkembangnya internet, maka kegiatan intelektual menggunakan format metadata atau secara sederhana data dari data. Format untuk keperluan Internet tersebut beraneka warna, misalnya Dublin Core,

Bibliografi

Almond, J. 1994. Ideas for information types and metadata attributes.

http://www.llnl.gov/liv_comp/metadata/papers/type_almond.ps

Burnett, Kathleen, Ng, Kwong Bor and Park, Soyeon. 1999. A comparison of the two traditions of metadata development. *Journal of the American Society for Information Science* 50 (13), 1209-1217

Caplan, P. 1995. You call it corn, we call it syntax-independent metadata for document-like objek. *The Public Access Computer Systems Review* 6 (4), 19-23.

<http://www.nlc.bnc.ca/documents/libraries/cataloging/caplan3.txt>

Chowdhury, G.G. 1999. *An introduction to modern information retrieval*. London: Library Association Publihsing

Dempsey, L., & Heery, R. 1997. A review of metadata: A survey of content resource description formats.

http://www.uklon.ac.kr/etadata/DESIRE/overview/rev_ti.html

Dempsey, L., & Weibel, S.L. 1996. The Warwick metadata workshop: A framework for the development of resource description. *D-Lib Magazine*, July/August 1996.

<http://www.dlib.org/dlib/Ju;y96/07Weibel.html>.

Desai, B. C. 1995. The semantic header and indexing and searching on the Internet.

<http://www.cs.concordia.ca/faculty/bcdesai/cindi-system-1.0tml>

Desai, B.C. Indexing and searching virtual libraries.

<http://www.cs.concordia.ca/faculty/bcdesai/forum95/forum95-bcd-indexing-html>

Deutsch, P., Emtage, A., Koster, M., & Stumpf, M. 1995. *Publishing information on the Internet with anonymous FTP*.

<http://info.webcrawler.com/mak/projects/iafa/iafa.text>

Freeman, G. and York, J. 1991. Client/server architecture promises radical changes, *CAUSE/EFFEC*, 14 (1).

<http://cause-www.colorado.edu/information-resources/ir-library/tex/cem9114.txt>

Gordano, R. 1994. The documentation of electronic text using text encoding initiative headers: An introduction. *Library Resources and Technical Services* 38 (4), 389-401

Gritton, Bruce, 1994. Metadata comments. http://www.iini.gov/liv_comp/metadata/papers/comments_gritton.html.

Hill, Linda et al. 1999. Collection metadata solutions for digital library applications.

Journal of the American Society for Information Science, 50 (13), 1169-1180

Hunter, C., and Springmeyer, R. 1994. Using metadata to create library guides for scientific analysis: A position paper on metadata by intelligent archive project at LLNL. Ftp://ftp.clearlake.ibm.com/pub/IEEE-Metadata/Archives_Workshop/Pos_Papers/Rspringmeyer

Jul, E. Internet cataloging project call for participation: building a catalog for Internet-accessible materials.

<http://www.oclc.org/oclc/man/catproj/catcall.html>.

Kim, Youngin, Norgard, Barbara and Chen, Aitao. 1999. "Using ordinary language to access metadata of diverse types of information resources: trade classifications and numeric data" dalam *Knowledge: creation, organization, and use: proceedings of the 62nd Annual Meeting of the American Society for Information Science*. pp:172-180 Medford, NJ: Information Today, Inc.

Madsen, M.S., Fogg, I. and Ruggles. C. 1994. Metadata systems: Integrative information technologies. *Libri*, 44 (3), 237-257

Ng, Kwong Bor, Park, Soyeon and Burnett, Kathleen. 1997. Control or management: A comparison of the two approaches for establishing metadata schemes in the digital environment. *ASIS 97: Digital collections: implications for users, funders, developers and maintainers: Proceedings of the 60th Annual Meeting of the American Society for Information Science*. p.:337-345. Medford NJ: Information Today Inc.

Rao, R. et al. 1995. Rich interaction in the digital library. *Communications of the ACM* 38 (4), 29-29

- Shelley, E.P. and Johnson, B.D. 1995. Metadata: concept and models. *Proceedings of the national conference on the management of geo-science information and data*, organized by Australian Mineral Foundation, Adelaide, Australia, 18-20 July 1995. p:1-5
- Shet, A.P. and Larson, J.A. 1990. Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Survey*, 22, 183-226
- Smith, T.R. 1996. The meta-information environment of digital library. *D-Lib Magazine*, July/August 1996.
<http://www.dlib.org/dlib/july96/new/07smitj.html>
- Smith, T.R., Gfeffner, S. and Gottsegen, J. 1996. A general framework for the meta-information and catalogs in digital libraries. Alexandria digital library public documents.
http://alexandria.sdc.ucsb.edu/public_documents/iee
- Strawman, A. and Bretherton, F. 1994. A reference model for metadata.
http://www.llnl.gov/liv_comp/metadata/papers/whitepaper-bretherton
- Strebel, D., Meeson, B. and Frithesen, J. 1994. Metadata standards and concepts for interdisciplinary scientific system – II. Position papers from IEEE Metadata Worksop (May 1994 – Washington D.C.)
http://clearlake.ibm.com/pub/IEEE=Metadata/Archives/Workshop/Pos_Papers/Donald_Strebel
- Sulistyo-Basuki. 2000. *Pengantar Ilmu Perpustakaan dan Informasi*. Diklat.
- Sutton, Stuart A et al. 1998. Networked information discovery and retrieval for educational materials on the Internet: metadata development, deployment, and evolution. *ASIS 98: Information access in the global information economy: Proceedings of the 61st Annual Meeting of the American Society for Information Science*. p.54-62. Medford NJ: Information Today Inc
- Taylor, Arlene. 1999. *The organization of information*. Englewood,CO: Libraries Unlimited, 1999
- Taylor, A.G. and Clemson, P. 1996. Access to networked documents: Catalog? Search engines? Both? OCLC Internet Cataloging Project Colloquium, position paper.
<http://www.oclc.org/oclc/man/colloq/taylor.html>
- Thornely, Jenny. 1998. The road to meta: the implementation of Dublin Core metadata in the State of Library of Queensland website. *The Australian Library Journal*, 47(1),74-82
- Triprabawa, Bagus. 1998. *Metadata, Dublin Core, Indomarc*. Presentasi di UPT Perpustakaan UI
- Weibel, S. 1995. Metadata: the foundations of resource description. *D-Lib magazine*.
<http://www.dlib.org/dlib/July95/07Weibel.html>
- Weibel, S. A proposed convention for embedding metadata in HTML.
<http://www.oclc.org/-weibel/html-meta.html>
- Weibel, S., Godby, J., Miller, E. and Daniel,R. 1995. The essential elements of network object description: OCLC/NCSA metadata workshop.
http://www.oclc.org:5046/oclc/research/conferences/metadata/dublin_core.html
- Woodward, J. 1996. Cataloguing and classifying information resources on the Internet. *Annual Review of Information Science and Technology*, 31,189-219

Xu, A. 1996. Accessing information on the Internet: feasibility study of USMARC format and AACR2.

<http://www.oclc.org/man/colloq/xu.html>

Younger, J.A. 1997. Resources description in a digital age. *Library Trends*, 45 (3),462-81

Zeng, Marcia Lei. 1999. Metadata elements for object description and representation: A case report from a digitized historical fashion collection project. *Journal of the American Society for Information Science*, 50 (13), 1193-1208

--oo0oo--