

**Penyebaran Ilmu dan Penerusan Tamadun
Melalui Teknologi Maklumat dan Komunikasi**

Tengku Mohd T. Sembok
Jabatan Sains Komputer
Universiti Islam Antarabangsa Malaysia
(tmts@iium.edu.my)

Pendahuluan

Saya ingin memulakan perbincangan dengan petikan kata-kata Dato' Hassan Ahmad, mantan Ketua Pengarah Dewan Bahasa dan Pustaka, Malaysia, dalam Sidang Ke-40 Majlis Bahasa Brunei Darussalam-Indonesia-Malaysia, 12-13 Mac 2001:

“...rancangan teknologi maklumat negara akan gagal jika ia tidak dikembangkan dalam bahasa Melayu, bahasa yang lebih layak daripada bahasa Inggeris untuk menjana kemajuan dan tamadun peribumi di negara ini”

Apakah tahap kebenaran kata-kata tersebut? Sebelum persoalan ini terjawab, beberapa konsep dan istilah perlu dikupas dan diterangkan perkaitannya. Konsep tersebut ialah teknologi maklumat, bahasa dan tamadun. Lantaran itu, bolehlah kita bincangkan kebenaran petikan kata-kata tersebut.

Tamadun

Peradaban atau tamadun ialah keadaan serta tingkat kemajuan pada kehidupan jasmani dan rohani sesuatu bangsa, masyarakat dan sebagainya (Kamus Dewan, 1993). Tamadun tidak kekal abadi tetapi melalui proses evolusi yang mempunyai fasa-fasa tertentu. Mengikut Quigley's *Evolution of Civilizations*, evolusi tamadun merangkumi fasa seperti *mixture, gestation, expansion, age of conflict, universal empire, decay, and invasion*. Sesuatu tamadun ditakrifkan dengan elemen sepunya seperti bahasa, sejarah, agama, budaya, institusi, dan jati diri warganya.

Juga terdapat ciri pemangkin tamadun yang diterima umum (Anwar, 2001, Sidang ke-40 Majlis Bahasa Brunei Darussalam-Indonesia-Malaysia), antaranya:

- Wujudnya sebuah masyarakat yang lebih kompleks dengan nilai dan norma tertentu;
- Masyarakat lebih berpengetahuan;
- Kemahiran dan pengkhususan;
- Bangunan yang unik atau skala besar;
- Teknologi metalurgi; dan
- Adanya bahasa yang berkembang pesat serta berfungsi.

Kedua-dua penulis, Quigley dan Anwar, menyebut unsur sepunya, iaitu *bahasa* dalam membina tamadun. Ini menunjukkan bagaimana pentingnya bahasa dalam membina dan meneruskan sesuatu tamadun. Hujah ini boleh digunakan untuk menyokong petikan kata-kata Dato' Hassan Ahmad:

“...rancangan teknologi maklumat negara akan gagal jika ia tidak dikembangkan dalam bahasa Melayu, bahasa yang lebih layak daripada bahasa Inggeris untuk menjana kemajuan dan tamadun peribumi di negara ini”

Alvin Toffler (Power Shift) pula mengutarakan bahawa infrastruktur elektronik yang menyokong masyarakat maklumat mengatasi unsur bangunan yang unik dan berskala besar dalam membina tamadun: *“Construction of electronic infrastructure of tomorrow’s super-symbolic society that dwarf anything that involve the building of the seven wonders of the worlds...”*

Kerajaan boleh bangun dan jatuh, pemerintah boleh datang dan pergi, tetapi tamadun mungkin kekal dalam kekalutan politik, sosial, ekonomi dan ideologi jika masih mempunyai warga yang aktif dalam menghayati serta menyebarkan faktor pemangkinnya. Huntington dalam Clash of Civilisations mengutarakan: *“All major scholars recognized the existence of a distinct Islamic civilization.... As a result, many distinct cultures or subcivilizations exist within Islam, including Arab, Turkic, Persian, and Malay.”* Mengikut beliau tamadun yang masih wujud lagi ialah China, Jepun, India, Islam dan Barat. Tamadun yang telah lenyap

pula ialah *Mesopotamian, Egyptian, Cretan, Classical, Byzantine, Middle American, Andean*.

Dua unsur pemangkin tamadun yang dapat kita rumuskan daripada sumber di atas ialah bahasa dan prasarana elektronik untuk meneruskan tamadun peribumi. Kedua-dua unsur ini akan dikupas secara mendalam.

Teknologi Maklumat dan Komunikasi

Sebelum rekaan mesin cetak, maklumat dan ilmu pengetahuan merupakan hak beberapa orang elit sahaja. Pada zaman itu sumber maklumat begitu terhad sekali dan sebarannya dilakukan melalui lisan dan tulisan tangan sahaja. Mesin cetak telah membolehkan maklumat dan ilmu pengetahuan dibuat salinan dengan mudah dan disebarkan dengan meluas. Mesin ini telah membawa paradigma baharu dalam penyebaran maklumat. Kini dengan ciptaan komputer, pada tahun 1940-an, satu revolusi baru telah bermula dalam cara manusia berurusan dengan maklumat dan ilmu. Revolusi ini telah membawa perubahan dalam semua aktiviti kehidupan manusia. Mengapakah ciptaan dua mesin ini telah mencetuskan paradigma baharu dan memberikan kesan yang begitu besar dalam kehidupan manusia?

Jelasnya kedua-dua mesin ini berurusan dengan maklumat. Mesin cetak memudahkan maklumat dan ilmu pengetahuan disalin dan disebarkan. Komputer pula ditakrifkan sebagai mesin pemprosesan maklumat. Komputer bukan hanya menolong dalam penyalinan dan penyebaran maklumat tetapi juga dalam pemprosesan maklumat yang meniru proses pemikiran manusia. Kefahaman yang lebih mendalam tentang pemikiran manusia akan melahirkan perisian komputer yang lebih baik dan cerdas. Di sebaliknya pula, dalam usaha melahirkan perisian yang lebih baik dan cerdas, kita akan lebih memahami proses pemikiran manusia.

Kemunculan rangkaian Internet telah memperkasakan lagi keupayaan komputer dalam pengendalian dan penyebaran maklumat. Komputer dan Internet telah memulakan proses pendemokrasian maklumat dan pengetahuan dengan meluas sekali dalam usaha mewujudkan sumber capaian dan sumber kreativiti kepada seluruh lapisan masyarakat. “Maklumat itu kuasa” merupakan slogan yang lazim diungkapkan oleh masyarakat yang digelar sebagai masyarakat maklumat atau terkini sebagai masyarakat pengetahuan dengan ekonomi pengetahuannya yang disebut sebagai k-ekonomi. Peralihan label daripada masyarakat maklumat kepada masyarakat pengetahuan menunjukkan kepekaan kepada hierarki pengetahuan atau ilmu. Konsep data, maklumat, pengetahuan, dan bijaksana mempunyai kaitan dengan peringkat pengetahuan yang berbeza.

Hierarki tahu bermula dengan data yang merupakan simbol dalam sesuatu medium penghantaran atau storan berpunca daripada sesuatu sumber. Data yang bermakna bagi penerima melahirkan maklumat baginya. Maklumat yang terkumpul dengan kepastian atau kebenaran yang terhasil daripada penggunaan maklumat melahirkan pengetahuan. Pengetahuan merupakan asas tindakan manusia dalam penyelesaian masalah. Dengan pengetahuan akan lahir kebijaksanaan. Bijaksana merupakan kebolehan mengambil tindakan yang paling baik berdasarkan pengetahuan yang terkumpul supaya melahirkan natijah yang paling baik. Kebijaksanaan juga melahirkan keupayaan kreativiti yang membolehkan seseorang melahirkan pengetahuan baharu.

Dalam masyarakat k-ekonomi, sebahagian besar pekerjaan merupakan kerja pengetahuan yang melibatkan aktiviti berdasarkan proses berfikir yang berintensifkan pengetahuan. Hasil kerja pengetahuan ialah penjanaan pengetahuan baharu. Oleh itu kerja pengetahuan ditakrifkan sebagai kerja yang melibatkan aktiviti mental yang kerap kali melahirkan pengetahuan baharu (Fisher & Fisher 1998). Tugas utama kerja ini ialah berfikir dan memerlukan kemahiran mental untuk berjaya. Untuk menghasilkan perkara baharu memerlukan bukan sahaja pengetahuan tetapi juga daya kreativiti.

Fokus utama penggunaan komputer semasa mula diciptakan pada tahun 1940-an, adalah dalam pengiraan data berangka. Fokus ini telah berubah kepada pengurusan data

(1960-an dan 1970-an), pengurusan maklumat (mulai 1980), dan akhir sekali kepada pengurusan pengetahuan (mulai 1995). Jadual 1 menunjukkan perubahan fokus penggunaan komputer serta kemahiran pekerja pada era yang berkaitan.

Teknologi terkini yang mencetuskan demokrasi menyeluruh dalam penjana, penyebaran, dan pengurusan maklumat ialah enjin gelintar dan Web 2.0 yang melahirkan Google, Wiki dan rangkaian sosial. Web 2.0 telah memusnahkreatifkan sempadan antara penjana dengan pengguna maklumat. Penjana dan pengguna maklumat tidak dapat dibezakan lagi dalam aplikasi seperti Wiki. Aplikasi seperti Facebook telah memperkasakan rangkaian sosial antara manusia dari setiap pelosok dunia. Warga Facebook akan mencecah satu bilion manusia tidak lama lagi dan akan merupakan *negara* teramai *penduduk*.

Jadual 1: Perubahan Fokus Penggunaan Komputer

Peringkat	1	2	3	4	5	6	7
Masa	1940-an	1960-an	1970-an	1980-an	1990-95	1995-2000	2001+
Fokus Teknikal	Algoritma Pengiraan Berangka	Pemusatan pengendalian Fail	Pengkomputeran Teragih	Pengkomputeran Peribadi	Rangkaian Setempat	Rangkaian Global dan Multimedia	Rangkaian Bergerak (Mobile)
Fokus Bisnes	Pengiraan Data Berangka	Pengurusan Data	Pengurusan Data	Pengurusan Maklumat	Pengurusan Maklumat	Pengurusan Pengetahuan dan Multimedia	Pengurusan Pengetahuan, Multimedia dan Bergerak Web 2.0; Wiki; Rangkaian Sosial; Web 3.0; Semantik; Awan(Cloud)
Kemahiran Pekerja	Rendah	Rendah	Rendah	Sederhana	Sederhana	Tinggi	Lebih Tinggi

Kini pemilik teknologi raksasa seperti Oracle, Microsoft dan Google memperoleh teknologi semantik dalam rangka peralihan daripada Web 2.0 kepada Web 3.0. Web 3.0 akan memperkasakan komputer dengan lebih luas dengan keupayaan memahami maklumat yang disimpan, diproses, dan disebar. Keupayaan ini membolehkan komputer secara cerdas

mengelintar, menyatukan, dan menjana pengetahuan baharu secara automatik. Syarat utama untuk mewujudkan keupayaan ini ialah memahami bahasa tabii.

Saya ingin mengambil kesempatan ini untuk membawa perbincangan kepada persoalan kecerdasan, bahasa, makna di peringkat falsafah, teori, model, dan aplikasi dari sudut teknologi maklumat amnya dan pemprosesan bahasa tabii khususnya.

Komputer dan Kecerdasan

Komputer banyak membantu manusia dalam mengendalikan maklumat dan pengetahuan. Dalam konteks ini, persoalan sama ada komputer mempunyai kecerdasan atau tidak masih menjadi tanda tanya. Filosof dan saintis telah lama membicarakan sama ada komputer mempunyai kecerdasan. Oleh sebab kesukaran untuk mendefinisikan kecerdasan secara objektif maka persoalan ini cuba dijawab menggunakan pendekatan behaviourisme (Sembok 2001a; 2001b). Sekiranya komputer mampu menunjukkan kelakuan yang ditentukan maka dapat dirumuskan bahawa komputer mempunyai kecerdasan. Berikut adalah antara perkara yang diutarakan oleh filosof dan saintis dalam isu ini.

Ujian

Alan Turing (1950) mengutarakan satu cadangan ujian terhadap kecerdasan komputer. Andaikan sebuah komputer diletakkan di dalam sebuah bilik dan kita boleh berhubung dengan komputer tersebut melalui sebuah terminal. Jika komputer dapat meyakinkan kita melalui dialog bahawa ia adalah manusia maka komputer tersebut lulus ujian Turing. Ini bermakna komputer mempunyai kecerdasan. Justeru, persoalannya ialah “adakah mesin dapat meyakinkan manusia bahawa ia adalah manusia?” Jika jawapannya “ya” maka komputer dibuktikan mempunyai kecerdasan!

Bilik Cina

John Searl (1980) mengutarakan satu senario yang dikenali sebagai Bilik Cina untuk menyanggahi ujian Turing. Ujarnya, sekiranya ada seorang yang tidak memahami bahasa Cina diletakkan di dalam sebuah bilik yang mempunyai dua pintu. Satu pintu digunakan untuk menerima masuk aksara Cina manakala satu pintu lagi untuk mengeluarkan aksara

Cina sebagai respons bagi aksara yang dimasukkan. Sekiranya orang ini hanya mahir dalam bahasa Inggeris dan dilengkapi dengan manual berbahasa Inggeris yang memberi arahan bagaimana untuk memberi respons bagi setiap aksara Cina yang diterimanya. Sekiranya orang ini bertindak seperti yang diarahkan oleh manual tersebut mungkin ia dapat meyakinkan orang di luar bahawa dia memahami bahasa Cina. Tetapi sebenarnya dia tidak memahami bahasa Cina. Begitu juga dapat dibicarakan tentang ujian Turing, kemungkinan komputer dapat meyakinkan manusia bahawa ia mempunyai kecerdasan, tetapi sebenarnya tidak.

Bilik Korea

Rapaport, pada tahun 1988, mengutarakan senario Bilik Korea untuk menyokong ujian Turing selepas digugat oleh senario Bilik Cina (Fetzer 1990). Andainya ada seorang profesor Korea yang mahir dalam bidang kajian penulisan Shakespeare. Profesor ini tidak memahami bahasa Inggeris tetapi kajian beliau berasaskan terjemahan yang baik dalam bahasa Korea. Komen dan kritikan profesor ini diterima baik oleh pakar dalam bidangnya dan dianggap sebagai pakar dalam pengajian Shakespeare. Hujahnya di sini, walaupun profesor tersebut tidak memahami bahasa Inggeris tetapi dia memahami sesuatu, khususnya bahasa Korea dan bahan penulisan Shakespeare dalam bahasa Korea yang membolehkan dia dianggap sebagai seorang intelektual dan pakar dalam pengajian Shakespeare.

Fenomena ini amat ketara dalam masyarakat Melayu yang tidak faham bahasa Arab. Tetapi orang Melayu boleh memahami agama Islam melalui bahasa Melayu. Kemungkinan ada orang Melayu yang tidak memahami bahasa Arab tetapi lebih alim dan mempunyai lebih ilmu tentang Islam berbanding dengan orang Arab yang memahami bahasa Arab. Cuma orang Islam seperti ini tidak diiktiraf boleh membuat fatwa dan digolongkan sebagai ulama kerana tidak mencukupi syarat yang ditetapkan oleh komuniti ulama. Tanpa memahami bahasa Arab mereka tidak boleh memahami al-Quran dan hadis melalui bahasa asalnya. Walau bagaimanapun mereka masih dapat memahami kandungan al-Quran dan hadis melalui bahasa ibundanya dan mengamalkan ajaran Islam. Orang Melayu begini mempunyai kecerdasan dan boleh masuk syurga. Bukankah begitu?

Bagi persoalan “adakah mesin dapat meyakinkan manusia bahawa ia adalah manusia?”, Rapoport merumuskan untuk berbuat demikian, komputer perlu memahami bahasa tabii. Memahami bahasa tabii merupakan syarat keperluan (*necessary condition*) untuk lulus ujian Turing dan membuktikan pemilikan kecerdasan. Rapoport juga percaya bahawa syarat ini juga syarat mencukupi (*sufficient condition*).

Bilik Melayu

Berdasarkan ujian Turing dan pendapat Rapoport, senario yang lebih mencabar boleh diutarakan. Ujian Turing menekankan dialog antara komputer dengan manusia untuk membuktikan kewujudan kecerdasan mesin. Kini dengan kemajuan yang pesat dalam teknologi bahasa dan perwakilan pengetahuan, sistem soal jawab atau dialog yang lebih canggih dapat dibangunkan. Kecerdasan sistem ini bergantung pada berapa banyak maklumat yang dimasukkan ke dalam komputer.

Senario yang lebih mencabar ialah yang menekankan kreativiti dan seni. Sekiranya sebuah komputer diletakkan di dalam sebuah bilik dan perhubungan dengan manusia dapat dilakukan melalui peranti input-output. Cabarannya ialah mampukah komputer menerima pantun dan mencipta pantun balas yang sesuai sehingga boleh meyakinkan manusia bahawa pantun yang dihasilkan ialah ciptaan manusia. Penghasilan pantun balas merupakan satu aktiviti kreatif dan seni manusia. Ujian ini bertujuan untuk menguji sama ada sebuah komputer boleh memahami pantun Melayu dengan kiasan bahasanya dan mencipta pantun balas yang sesuai, bererti dan indah dari aspek bahasa. Bilik Melayu boleh dikatakan lebih mencabar daripada Bilik ujian Turing kerana pantun mempunyai makna yang tersurat dan tersirat serta seni bahasa. Pantun balas pula perlu mempunyai ciri-ciri yang sama.

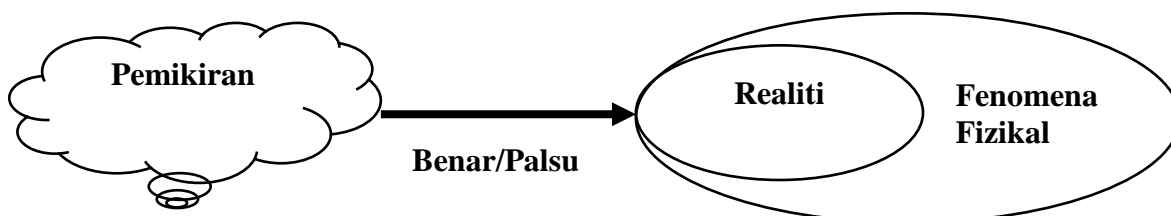
Kecerdasan dan Bahasa

Rapoport telah mensyaratkan penguasaan bahasa sebagai keperluan kepada pemilikan kecerdasan. Karl Popper pula mengutarakan tentang bahasa seperti yang berikut “... *the human language, the first and most important product of the human mind...*” (Popper 1999). Melalui bahasa manusia dapat mencerap realiti, menstrukturkan minda, memikirkan

masalah, menjanakan idea, mengungkapkan makna, dan memberi kesan kepada realiti. Manusia menggunakan bahasa dalam mengembara mencari kepastian dan akhirnya berjumpa dengan kebenaran.

Ahli sains dan falsafah seperti Bertrand Russell, Heidegger, Hume, dan Wittgenstein mengutarakan beberapa teori bahasa dan makna (Magee 1987). Wittgenstein dalam karya agungnya, *Tractatus Logico-Philosophicus*, mengutarakan teori gambaran makna (*the picture theory of meaning*) yang mengatakan bahawa bahasa adalah untuk menggambarkan atau mewakili realiti. Ayat mengandungi nama-nama yang berpadanan dengan objek di dunia. Susunan nama dalam ayat pula berpadanan dengan susunan objek di dunia. Maka kita boleh membaca struktur realiti daripada struktur ayat dalam sesuatu bahasa. Justeru struktur realiti yang menentukan struktur bahasa. Ini berkait rapat dengan teori padanan kebenaran (*correspondence theory of truth*) yang dipelopori oleh Karl Popper (1977). Teori ini menyatakan bahawa pemikiran seseorang adalah benar sekiranya ia sepadan dengan realiti. Realitilah yang membenarkan pemikiran (lihat Rajah 1). Sifat teori ini adalah pasti, deterministik, boleh ramal, dan menolak rujukan sendiri. Jawapan yang diberikan kepada soalan ialah benar atau palsu. Pandangan ini melahirkan ontologi dan taksonomi dalam menyusun hubung kait antara nama yang wujud secara amnya dan dalam sesuatu bidang khususnya.

Teori ini digunakan sebagai asas pembangunan model untuk melakukan eksperimen mencari kebenaran dalam penyelidikan. Apa yang dimodelkan dalam minda boleh dibuktikan benar atau palsu melalui eksperimen dan pencerapan. Contohnya, model yang menggambarkan dunia sebagai pusat sistem suria telah dibuktikan palsu oleh cerapan saintifik yang menunjukkan bahawa matahari ialah pusat sebenarnya. Perubahan model yang asas seperti ini boleh dikatakan sebagai satu anjakan paradigma yang membawa banyak perubahan ilmu yang berdasarkan model lama.



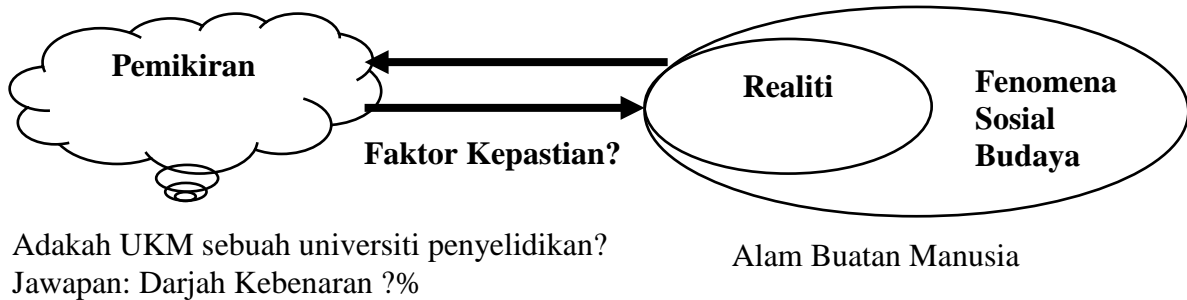
Adakah Universiti Kebangsaan Malaysia
terletak di Bangi?
Jawapan: Benar

Alam Tabii

Rajah 1: Model Teori Padanan Kebenaran

Wittgenstein membaharui pandangannya tentang bahasa dalam buku keduanya, *Philosophical Investigation*, yang mengutarakan konsep alat bagi makna: maksudnya makna sesuatu perkataan ialah sebagaimana ia digunakan dalam bahasa. Maka timbullah konsep keserupaan keluarga (*family resemblance*) dalam penggunaan perkataan. Ini memberi asas dalam kajian bahasa melalui korpus. Pandangan baharu ini mencetuskan idea bahawa struktur bahasa kita yang menentukan bagaimana kita berfikir tentang dunia sebenar. Kita tidak boleh membincangkan tentang dunia dan tidak boleh berfikir tentang dunia tanpa alat konseptual untuk tujuan tersebut. Alat ini wujud dalam bahasa. Justeru, struktur bahasa menentukan struktur realiti. Realiti hanyalah apa-apa yang kita fikirkan, itulah yang dinyatakan oleh falsafah idealisme. Wacana membincangkan kenyataan realiti semasa merupakan salah satu daripada pelbagai wacana yang mungkin. Membincangkan alam utopia juga satu wacana. Idea ini mempunyai kaitan dengan hipotesis refleksiviti (*reflexivity hypothesis*) yang diutarakan oleh Soros (Soros 1998) (lihat Rajah 2). Soros menggunakan idea ini untuk membentuk realiti masa depan berdasarkan realiti sekarang dan realiti yang diidamkan, mungkin dalam perniagaannya yang melibatkan penukaran mata wang. Maka natijahnya boleh dikatakan bahawa realiti boleh menstrukturkan fikiran dan fikiran boleh menstrukturkan realiti. Hipotesis refleksiviti mempunyai ciri-ciri tak deterministik, tak pasti, tak boleh ramal, dan menerima rujukan sendiri. Jawapan yang diberikan kepada soalan ialah dalam darjah kepastian (Sembok 1987).

Penggunaan strategi hipotesis refleksiviti perlu dibatasi dengan nilai yang murni dalam manipulasian fenomena sosial untuk mencapai matlamat. Rasionaliti tak berbatas dalam kejuruteraan sosial boleh membawa kerosakan kepada masyarakat. Batasan yang paling unggul ialah yang berdasarkan *sunnatullah*.



Rajah 2: Hipotesis Refleksiviti

Bahasa dan Budaya

Ahli sosiologi bahasa pula memberikan tiga aras dalam pengkajian bahasa dan budaya, iaitu aras sosial (*sinkronik*), sejarah (*diakronik*), dan khayalan (*imaginasi*) (Kramsch 1998). Komuniti wacana dicirikan bukan sahaja dengan fakta dan artifak, tetapi juga dengan impian, khayalan yang tercapai dan tak tercapai. Seperti slogan *Malaysia Boleh*, dan juga unsur-unsur yang *tak boleh* dalam masyarakat Malaysia kini. Mereka yang berpengaruh akan mencorakkan masyarakat, menentukan nilai dan kepercayaan yang dianuti oleh komuniti, peristiwa sejarah yang perlu diraikan, dan masa depan yang perlu dikhayalkan. Ini memberi makna bahawa bahasa bukan satu kod yang bebas budaya, yang tidak berkaitan dengan cara berfikir dan kelakuan ahli komuniti.

Budaya boleh ditakrifkan sebagai kriteria keahlian dalam satu komuniti wacana yang berkongsi ruang sosial dan sejarah sepunya serta impian sepunya. Walaupun ahlinya telah meninggalkan komuniti tersebut, mereka masih lagi memegang sistem piawaiannya sepunya dalam persepsi, kepercayaan, penilaian, tindakan, dan tutur bahasa. Piawaiannya ini dikenali sebagai *budaya*. Pandangan budaya yang berbeza dan hubungannya dengan bahasa menimbulkan persoalan yang fundamental: sejauh manakah tasawur dan aktiviti mental seorang ahli komuniti dilenturkan atau bersandarkan bahasa yang digunakan? Hipotesis Sapir-Whorf menyatakan bahawa struktur bahasa yang digunakan seharian mempengaruhi cara seseorang berfikir dan berkelakuan. Berpegang pada prinsip ini maka satu bahasa komputer Ungkapan Komputer Melayu Pertama (UKM1) telah dibangunkan dan digunakan untuk mengajarkan kursus pengatur cara pada tahun 1984 (Sembok 1984). Contoh atur cara

UKM1 diberikan dalam Rajah 3. Pengajaran pengaturcaraan boleh dilaksanakan dengan menggunakan bahasa Melayu sepenuhnya tanpa alih kod (*code switching*) dengan adanya UKM1 kerana bahasa pengaturcaraan yang lain adalah dalam bahasa asing.

```
atur cara faktoran (output);
(* atur cara mengira nilai faktoran daripada 1 hingga 25 *)
Pemboleh ubah i, hasil darab : integer;
mula tulis baris('      ', 'N', '      N!');
    hasil darab:=1;
    bagi i:= 1 hingga 25 buat
        mula
            hasil darab := hasil darab*i;
            tulis baris('      ', i, '      ', hasil darab);
        tamat
tamat
```

Rajah 3: Atur cara dalam bahasa UKM1

Melalui bahasa manusia memahami dunia dan memodelkan dunia untuk menyelesaikan masalah atau memodelkan dunia impian sebagai sasaran. Dalam teknologi maklumat pemodelan dunia realiti untuk menyelesaikan masalah atau memodelkan dunia hipotektikal untuk produk baharu dengan mengambil kira nilai dan budaya merupakan kerja pengetahuan yang perlu dihadapi untuk *survival* dalam alaf baru ini. Pendekatan hipotesis reflektiviti mungkin dapat menolong dalam proses pemodelan ini. Sebelum wujudnya e-mel, Internet, dan enjin gelintar, pencipta produk ini telah membayangkan senario budaya dunia siber yang menggunakan produk ciptaan tersebut.

Fitrah Kebolehan Berbahasa

Dalam tempoh dua dekad yang lepas, kepercayaan bahawa kebolehan berbahasa mempunyai kaitan dengan komponen genetik telah mempengaruhi penyelidik dalam bidang bahasa. Ini bermakna kebolehan berbahasa merupakan fitrah kejadian manusia. Manusia boleh berbicara dan monyet tidak boleh adalah kerana pembezaan genetik. Penyelidik percaya ada

dua cara bagaimana kebolehan berbahasa boleh dikaitkan dengan fitrah semula jadi. Pertama ialah manusia lebih cerdas daripada monyet maka kebolehan berbahasa merupakan hasilnya. Kedua ialah manusia mempunyai organ bahasa dalam otaknya yang telah diset tanpa rangsangan atau stimulus luaran (Tulving 1997). Ini umpamanya sebahagian rangkaian neural otak manusia telah diset untuk menerima bahasa. Kajian loka sian dalam fungsi otak menguatkan kepercayaan fitrah dengan penemuan bahagian otak yang dinamakan bahagian Broca dan bahagian Wernicke. Kerosakan pada bahagian Broca mengakibatkan kepayahan menguasai nahu, dan kerosakan pada bahagian Wernicke pula mengakibatkan kebolehan mengeluarkan ayat yang betul dari segi nahu tetapi tidak bermakna (Thompson 1997).

Noam Chomsky dan pengikutnya merupakan penyelidik yang percaya kepada kesemulajadian berbahasa pada manusia. Mereka telah memberikan sumbangan yang besar dalam perkembangan linguistik, antara lain ialah tiga pengamatan seperti yang berikut (Kramsch 1998):

- i. Setiap bahasa tabii, seperti bahasa Melayu dan Inggeris, mempunyai satu set petua yang terhingga. Dengan menggunakan petua ini, semua ayat yang sah dari segi nahu dalam sesuatu bahasa dapat dijanakan. Senarai petua ini dikenali sebagai *nahu generatif*. Pendapat ini memberi nafas dan menghidupkan bidang pengkomputeran linguistik.
- ii. Kanak-kanak, tak kira asal genetik mereka, mampu menguasai mana-mana bahasa tabii. Ini juga berlaku kepada orang dewasa dengan syarat mereka telah menguasai bahasa ibunda mereka terlebih dahulu. Ini bermakna wujudnya kebolehan umum untuk menguasai mana-mana nahu generatif atau kebolehan menguasai *nahu universal*.
- iii. Nahu universal mempunyai komponen semula jadi yang kukuh.

Pengamatan ini telah membawa perkembangan yang pesat dalam bidang pemprosesan dan teknologi bahasa. Kewujudan nahu generatif memberi keupayaan kepada komputer memproses, memahami, dan menjana bahasa tabii. Ini boleh diibaratkan seperti memori komputer telah diset untuk memahami bahasa sebagai mana otak manusia. Implikasi kepada

manusia pula, ia memberi harapan bahawa sesiapa sahaja boleh menguasai bahasa asing. Orang Melayu boleh menguasai bahasa Inggeris secara tabii.

Senario dan teori yang dikemukakan memberikan implikasi yang amat positif kepada bahasa Melayu untuk menjadi bahasa ilmu yang unggul dan keupayaan orang Melayu menguasai apa-apa sahaja bahasa asing yang diinginkan. Maka yang menentukan kejayaan bergantung pada kehendak dan usaha yang gigih.

Pemprosesan Bahasa Tabii: Satu Kecerdasan

Seperti diutarakan oleh Chomsky bahawa setiap bahasa tabii mempunyai satu set petua yang terhingga yang boleh digunakan untuk menjana semua ayat yang sah dari segi nahunya. Malah petua tersebut boleh digunakan juga untuk menyemak sintaks ayat dan memahami semantiknya bagi sesuatu bahasa. Bidang pemprosesan bahasa tabii memberi tumpuan pada pengumpulan petua-petua ini supaya komputer dapat memproses bahasa manusia serta memahaminya. Lebih-lebih lagi dengan kemunculan Internet dan Web, telah berlaku satu ledakan khazanah ilmu dalam bentuk digital dan khazanah ini dinukilkan dalam bahasa tabii. Tanpa bantuan komputer, manusia sukar untuk memproses, mengorganisasikan, menyimpan, dan mencapai kembali ilmu tersebut.

Petua yang penting dalam pemprosesan bahasa tabii ialah petua fonologi, morfologi, sintaks, semantik, pragmatik, serta petua untuk mewakili pengetahuan am atau dikenali sebagai pengetahuan dunia seperti ditunjukkan dalam Jadual 2. Di samping mempunyai petua yang berkaitan dengan bahasa, petua untuk melakukan inferens juga diperlukan dalam sesebuah sistem aplikasi pemprosesan bahasa tabii. Petua ini membolehkan sistem melaksanakan deduksi atau induksi dalam menyelesaikan sesuatu masalah. Kebolehan melaksanakan inferens memberikan ciri kecerdasan kepada komputer bagi meniru satu sifat kemanusiaan.

Jadual 2: Petua dan Kecerdasan dalam Pemprosesan Bahasa Tabii

Hierarki Petua	Fungsi	Contoh
Inferens	Kecerdasan membuat deduksi dan induksi untuk	Aminah telah bercerai mati atau hidup dengan suami

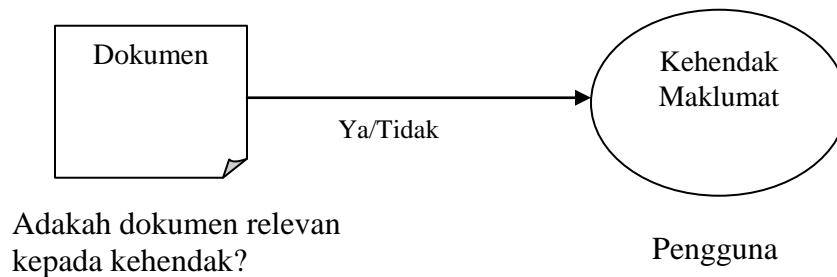
	menyelesaikan masalah	terdahulu.
Pengetahuan Dunia	Pengetahuan am yang membantu makna dan maksud	Wanita boleh berkahwin baru sekiranya tanpa suami lagi. Suami ialah seorang lelaki.
Pragmatik	Maksud perkataan dan ayat dalam konteks	Suami baru => bukan suami pertama.
Semantik	Makna perkataan dan ayat	Suami(Amin, Aminah)
Sintaks	Bagaimana perkataan membina struktur ayat	Amin(nama) ialah(sendi kata) suami(nama) baru(sifat) aminah(nama).
Morfologi	Struktur dan pembentukan perkataan	Amin, Amin ah
Fonologi	Bunyi sebutan perkataan	Aminah = “A” “mi” “nah”

Contoh Teknologi Peribumi dalam Capaian Maklumat

Kembali kepada petikan kata-kata Dato’ Hassan Ahmad, sekiranya rancangan teknologi maklumat negara dikembangkan dalam bahasa Melayu, maka akan muncullah teknologi berkaitan dengan bahasa Melayu. Di sini, kita utarakan teknologi capaian maklumat untuk memproses kandungan yang ditulis dalam bahasa Melayu. Enjin gelintar memerlukan indeks dalam bahasa Melayu untuk menyimpan dan mencapai kembali dokumen dalam bahasa Melayu.

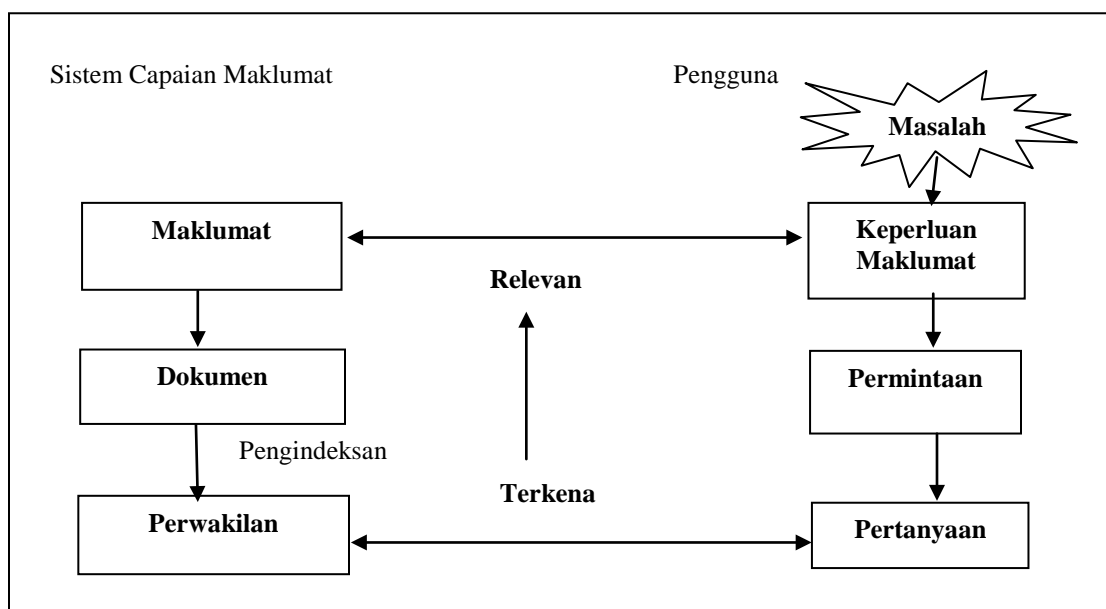
Segala ilmu, pengetahuan dan maklumat diungkapkan, disimpan, disebar, dan dibincangkan menggunakan bahasa tabii. Kini pengetahuan, dan maklumat telah dikendalikan dalam bentuk digital sejak mula ciptaan komputer. Kemunculan Internet dan sesawang merentas dunia (*world wide web*) membolehkan laman sesawang dibangunkan dan dirangkaikan di seluruh dunia. Pelbagai maklumat dan dokumen telah dimuatkan dalam laman-laman ini untuk dicapai dan dirujuk oleh pelayar Internet. Tetapi malangnya

maklumat yang dikehendaki tidak semuanya boleh dicapai oleh enjin gelintar. Proses menggelintar dan menentukan dokumen mana yang relevan bukanlah usaha remeh. Proses ini melibatkan memahami kehendak pengguna dan isi kandungan dokumen yang ditulis dalam bahasa tabii. Untuk menentukan sama ada sesuatu dokumen relevan atau tidak bagi pengguna merupakan satu proses intelektual (lihat Rajah 4). Komputer hanya dapat menganggarkan kerelevanan sesuatu dokumen kepada pengguna. Kerelevanan sebenar hanya dapat ditentukan oleh pengguna sendiri (Sembok 1998).



Rajah 4: Konsep Kerelevanan

Capaian maklumat ditakrifkan sebagai bidang kajian bagaimana hendak menentukan dan mencapai maklumat yang relevan dengan keperluan daripada korpus maklumat yang terstor (van Rijsbergen 1979). Andainya terdapat satu stor maklumat, seperti sesawang (*web*), yang besar yang mengandungi dokumen dalam pelbagai bidang. Pengguna stor ini mempunyai masalah untuk diselesaikan dan memerlukan maklumat yang relevan dengan masalah berkenaan. Pengguna perlu mengungkapkan keperluan maklumat ini sebagai permintaan (*request*) dalam satu bentuk kepada stor ini. Permintaan ini diterjemahkan sebagai pertanyaan (*query*) yang merupakan satu perwakilan yang ditetapkan oleh sesebuah sistem capaian maklumat. Sistem capaian maklumat pula hanya menyimpan perwakilan dalam dokumen yang mengandungi maklumat. Apabila sistem menerima pertanyaan daripada pengguna, perwakilan dokumen akan digunakan untuk menentukan kebarangkalian sama ada dokumen berkenaan relevan dengan keperluan maklumat. Semua dokumen yang mencapai tahap kebarangkalian tertentu akan diberikan kepada pengguna dalam bentuk senarai *terkena* (*hits list*). Pengguna akan melihat dokumen yang tersenarai dan menentukan sama ada relevan atau tidak. Rajah 5 memberikan gambaran bagi menjelaskan perkara yang dibincangkan (Mizzaro 1997).



Rajah 5: Perwakilan dan Kerelevanan Dokumen

Perwakilan Dokumen

Kandungan utama sesebuah dokumen ialah perkataan serta ditokok tambah dengan rajah dan jadual. Perkataan digunakan untuk mewakili dokumen bagi tujuan pengindeksan supaya dokumen senang distorkan dan dicapai kembali. Pemilihan perkataan sebagai indeks merupakan satu bidang kajian tersendiri supaya perkataan yang dipilih dapat menjadi pemerihal yang tepat kepada dokumen. Ada juga perkataan yang tidak sesuai digunakan dalam pengindeksan kerana perkataan ini tidak boleh membezakan satu dokumen dengan yang lain. Perkataan ini dikenali sebagai kata henti (*stop words*).

Dokumen dapat diwakilkan dengan pelbagai cara mengikut aras pemprosesan yang dilakukan. Teori aras pemprosesan maklumat (*level of processing theory*) menyatakan bahawa wujudnya pelbagai cara untuk memproses dan mengkodkan maklumat (Gagne 1993). Aras pemprosesan yang digunakan akan menentukan kualiti perwakilan maklumat dalam storan komputer. Aras pemprosesan maklumat dalam sistem capaian maklumat boleh dikelaskan seperti yang berikut: pengkodan rentetan, morfologi, sintaksis, dan semantik (Fagan 1987; Smeaton 1988; Sembok 1989).

Model capaian maklumat yang dipilih akan menentukan aras-pemprosesan yang

digunakan. Contohnya, model Boolean dan model ruang vektor menggunakan aras rentetan dan morfologi untuk mengindeks dan mencapai dokumen. Model ini menggunakan rentetan kata kunci yang tak bersandaran sebagai perwakilan dokumen. Model linguistik menggunakan teknik pemrosesan bahasa tabii dan kecerdasan buatan untuk mewakili dokumen dalam bentuk perwakilan berasaskan pengetahuan (*knowledge based representation*).

Persoalan Persamaan Perkataan

Persoalannya timbul dalam situasi bagi pertanyaan yang mengandungi perkataan “makanan” dan dokumen hanya mengandungi perkataan “pemakanan”. Adakah perkataan ini boleh dianggap sama dan dokumen ini mempunyai nilai kepastian kerelevanan kerana perkongsian kata akar? Vektor pertanyaan K dan vektor dokumen D di bawah menjelaskan persoalan yang dikemukakan.

$$K = (\dots; \text{makanan}, 1; \text{pemakanan}, 0; \dots)$$

$$D = (\dots; \text{makanan}, 0; \text{pemakanan}, 0.9; \dots)$$

$$\text{Kepastian}(D, K) = 0 \text{ atau } > 0?$$

Penyelidik tentang *makanan* mungkin juga mahu membaca dokumen yang mengandungi perkataan-perkataan *makan*, *dimakan*, *termakan*, *makanannya*, *pemakanan*, dan perkataan-perkataan lain yang mempunyai kata akar yang sama dengan harapan dokumen-dokumen berkenaan membincangkan perkara yang berkaitan.

Algoritma Cantasan Melayu

Satu cara penyelesaian kepada persoalan persamaan perkataan ialah menggunakan algoritma pengakar atau cantasan (*stemming*) untuk menggabungkan (*conflate*) perkataan yang berkongsi kata akar atau kata dasar. Algoritma ini telah wujud untuk pelbagai bahasa, seperti

algoritma Porter (1980) untuk bahasa Inggeris, Savoy (1993) untuk bahasa Perancis, Al-Omari (1994) dan Abu Ata (2001) untuk bahasa Arab, dan Popovic & Willett (1992) untuk bahasa Slovene. Pengindeksan maklumat menggunakan kaedah gabungan telah dibuktikan boleh memperbaiki keberkesanan capaian (Harman 1991; Abu Ata 2001; Ahmad 1996).

Algoritma Othman-Sembok

Bagi bahasa Melayu algoritma cantasan Othman & Sembok (1993) merupakan antara yang pertama. Dalam algoritma ini jenis imbuhan bahasa Melayu dibahagikan kepada empat jenis, iaitu apitan, awalan, akhiran dan sisipan seperti dalam Jadual 3. Petua ejaan yang melibatkan kata akar yang bermula dengan huruf f, p, k, s, dan t apabila digunakan dengan imbuhan mem, pem, meng, peng, meny, peny, men, dan pen juga diambil kira oleh sistem seperti contoh berikut:

- mem + fikir = memikir
- mem + pukul = memukul
- meng + karang = mengarang
- men + tuai = menuai
- meny + simpang = menyimpang.

Jadual 3: Jenis Petua Imbuhan

Jenis Petua Imbuhan	Contoh Petua	Contoh Perkataan
Apitan (ap)	ke+an ke+annya peng+an	kecerdasan keutamaannya pengertian
Awalan (aw)	be+ bel+ ber+	bekerja belajar beradik
Akhiran (ak)	+an +wi +at	makanan duniawi muslimat
Sisipan (sp)	+in+ +em+ +er+	sinambungan gementar gerigi

Algoritma Othman-Sembok diuji dengan sepuluh surah al-Quran terjemahan bahasa Melayu yang mengandungi 736 perkataan unik. Sebanyak 121 petua imbuhan digunakan. Hasil yang diperoleh boleh diukur melalui bilangan ralat yang dilahirkan, iaitu sebanyak 96 kesilapan (lihat Jadual 4). Sembok et al. (1995; 1999) dan Abu Ata et al. (1997) telah mengimplementasikan algoritma ini untuk mengindeks dokumen dengan menggunakan kaedah cincang sebagai capaian langsung kepada indeks untuk membangunkan sistem capaian dokumen multilingual dan sistem capaian istilah.

Jadual 4: Ralat daripada Eksperimen Algoritma Tertib Aplikasi Petua

Perkataan	Akar Sebenar	Hasil Cantasan	Jenis Ralat
kurangkan memakan sebabnya berduri	kurang makan sebab duri	rang mak bab dur	lebih pantas lebih pantas lebih pantas lebih pantas
peringatan	ingat	peringat	kurang pantas
sepuh perangan berbuah sebelah sukai	puh perang buah belah suka	sepuh perangan berbuah sebelah sukai	tak berubah tak berubah tak berubah tak berubah Tak berubah
mengandung	kandung	gandung	ejaan
berilah berikanlah beriman mencari keburukan melihat sekali melengah memulai kesenangan menurunkan	beri beri iman cari buruk lihat kali lengah mula senang turun	ilah ikan rim pencar keburu pelih sekal meleng pulai tangan penurun	lain-lain lain-lain lain-lain lain-lain lain-lain lain-lain lain-lain lain-lain lain-lain lain-lain

Rumusan

Komponen sistem teknologi maklumat mempunyai empat unsur utama, iaitu komputer, komunikasi, kandungan dan komuniti. Komputer dan rangkaian komunikasi merupakan prasarana yang penting dalam penjanaaan dan penyebaran ilmu pengetahuan yang lebih penting daripada bangunan raksasa dalam pemangkinan tamadun. Perkakasan komputer dan alat komunikasi boleh dikatakan bebas budaya dan bahasa. Kandungan pula tidak bebas daripada sosiobudaya dan bahasa. Kandungan merupakan inti pati ketamadunan yang perlu disimpan, diproses dan disebarkan kepada warga komuniti lokal serta global. Sekiranya kandungan disimpan dalam bahasa Melayu untuk menyebarkan ilmu dan melatari tamadun maka teknologi pemprosesan bahasa Melayu perlu diwujudkan.

Wiki, e-mel dan rangkaian sosial melibatkan berbilion ahli komuniti dari setiap pelosok dunia. Maka setiap kandungan yang dijana untuk tatapan pengguna Internet akan memberi kesan kepada komuniti pengguna secara positif atau negatif. Justeru, petikan kata-kata Dato' Hassan Ahmad ada kebenarannya.

Beberapa teori dan hipotesis seperti ujian Turing, Bilik Cina, Bilik Korea, Bilik Melayu, hipotesis Sapir-Whorf, teori gambaran makna Wittgentein, teori padanan kebenaran Karl Popper, dan hipotesis reflektiviti Soros dikupas untuk memahami ciri dan isu yang berkaitan dengan pemikiran, kecerdasan, bahasa, nilai, dan budaya.

Keseluruhannya falsafah, model, reka bentuk dalam pemikiran, kecerdasan, dan bahasa telah dibincangkan dari sudut teknologi maklumat secara amnya dan aplikasi dalam capaian maklumat secara khususnya berdasarkan rangka pengetahuan holistik. Usaha utama yang patut ditumpukan demi kemajuan teknologi bahasa Melayu ialah pembentukan leksikografi elektronik lengkap bagi perkataan Melayu. Leksikografi ini perlu mengandungi maklumat asas bagi setiap perkataan dari segi fonem, morfologi, sintaksis, semantik, dan pragmatik. Nahu pengkomputeran untuk bahasa Melayu juga perlu dilengkapi supaya pemproses, seperti huraian, penterjemahan, dan pemahaman ayat dapat dilakukan secara automatik. Dengan adanya bahan elektronik asas ini akan membuka pintu seluas-luasnya untuk penyelidikan teknologi dalam bahasa Melayu.