

COLEÇÃO PALAVRA-CHAVE

Roberto Barsotti

A
INFORMÁTICA
NA
BIBLIOTECONOMIA
E NA
DOCUMENTAÇÃO

editora polis
apb

ROBERTO BARSOTTI

A INFORMÁTICA NA BIBLIOTECONOMIA E NA DOCUMENTAÇÃO

**editora polis
associação paulista de bibliotecários
1990**

Copyright © 1990 do autor
Capa: Sandra Regina Sumye Kato

BARSOTTI, Roberto. *A informática na biblioteconomia e na*
B282i *documentação*. São Paulo : Polis, APB, 1990. (Coleção Pa-
lavra-chave, 2)
127 p.

1. Bibliotecas — Automação I. Título

CDD — (19a)020.285
CDU — 02.681

Índice para catálogo sistemático:

1. Informática : Biblioteconomia
2. Biblioteconomia : Informática

ISBN: 85-7228-001-4

Direitos reservados pela
LIVRARIA E EDITORA POLIS LTDA.
Rua Caramuru, 1196 — Saúde — 04138 — São Paulo — SP
Tel.: (011)275-7586

SUMÁRIO

Introdução	1
PREFÁCIO	11
1. A MÃE	13
2. O PAI	15
3. O FILHO	17
4. O AMOR	19
5. A VIDA	21
6. A MORTE	23
7. A ESPERANÇA	25
8. A FÉ	27
9. A CARIDADE	29
10. A PAZ	31
11. A JUSTIÇA	33
12. A VERDADE	35
13. A LIBERDADE	37
14. A IGUALDADE	39
15. A HUMANIDADE	41
16. A CIDADANIA	43
17. A RESPONSABILIDADE	45
18. A RESPEITO	47
19. A TOLERÂNCIA	49
20. A COOPERAÇÃO	51
21. A PARTICIPAÇÃO	53
22. A TRANSFORMAÇÃO	55
23. A CRIATIVIDADE	57
24. A INICIATIVA	59
25. A PERSEVERANÇA	61
26. A RESILIENTE	63
27. A OPORTUNIDADE	65
28. A INOVAÇÃO	67
29. A SUSTENTABILIDADE	69
30. A ÉTICA	71
31. A CONSCIENTIZAÇÃO	73
32. A PARTICIPAÇÃO	75
33. A TRANSFORMAÇÃO	77
34. A CRIATIVIDADE	79
35. A INICIATIVA	81
36. A PERSEVERANÇA	83
37. A RESILIENTE	85
38. A OPORTUNIDADE	87
39. A INOVAÇÃO	89
40. A SUSTENTABILIDADE	91
41. A ÉTICA	93
42. A CONSCIENTIZAÇÃO	95
43. A PARTICIPAÇÃO	97
44. A TRANSFORMAÇÃO	99
45. A CRIATIVIDADE	101
46. A INICIATIVA	103
47. A PERSEVERANÇA	105
48. A RESILIENTE	107
49. A OPORTUNIDADE	109
50. A INOVAÇÃO	111
51. A SUSTENTABILIDADE	113
52. A ÉTICA	115
53. A CONSCIENTIZAÇÃO	117
54. A PARTICIPAÇÃO	119
55. A TRANSFORMAÇÃO	121
56. A CRIATIVIDADE	123
57. A INICIATIVA	125
58. A PERSEVERANÇA	127
59. A RESILIENTE	129
60. A OPORTUNIDADE	131
61. A INOVAÇÃO	133
62. A SUSTENTABILIDADE	135
63. A ÉTICA	137
64. A CONSCIENTIZAÇÃO	139
65. A PARTICIPAÇÃO	141
66. A TRANSFORMAÇÃO	143
67. A CRIATIVIDADE	145
68. A INICIATIVA	147
69. A PERSEVERANÇA	149
70. A RESILIENTE	151
71. A OPORTUNIDADE	153
72. A INOVAÇÃO	155
73. A SUSTENTABILIDADE	157
74. A ÉTICA	159
75. A CONSCIENTIZAÇÃO	161
76. A PARTICIPAÇÃO	163
77. A TRANSFORMAÇÃO	165
78. A CRIATIVIDADE	167
79. A INICIATIVA	169
80. A PERSEVERANÇA	171
81. A RESILIENTE	173
82. A OPORTUNIDADE	175
83. A INOVAÇÃO	177
84. A SUSTENTABILIDADE	179
85. A ÉTICA	181
86. A CONSCIENTIZAÇÃO	183
87. A PARTICIPAÇÃO	185
88. A TRANSFORMAÇÃO	187
89. A CRIATIVIDADE	189
90. A INICIATIVA	191
91. A PERSEVERANÇA	193
92. A RESILIENTE	195
93. A OPORTUNIDADE	197
94. A INOVAÇÃO	199
95. A SUSTENTABILIDADE	201
96. A ÉTICA	203
97. A CONSCIENTIZAÇÃO	205
98. A PARTICIPAÇÃO	207
99. A TRANSFORMAÇÃO	209
100. A CRIATIVIDADE	211

*À minha mãe,
que me ensinou a amar.*

*Ao meu pai,
que me ensinou a ser honrado.*

SUMÁRIO

1. INTRODUÇÃO	9
2. INFORMÁTICA TEÓRICA	11
2.1. O computador	11
2.1.1. Breve histórico	11
2.1.2. Gerações	14
2.1.3. Tipos	16
2.1.4. Componentes	16
2.1.4.1. Veículos de dados e/ou informações .	18
2.1.4.2. Unidade Central de Processamento .	23
2.2. Sistemas de numeração .:.....	25
2.3. Representação das informações no computador	28
2.4. Principais códigos utilizados pelos computadores ...	30
2.5. Organização de arquivos	36
2.6. Fluxogramas	41
2.7. Programação e linguagens	49
2.7.1. Software aplicativo e software básico	53
2.8. Teleprocessamento	54
2.8.1. Características do Teleprocessamento	56
2.8.2. Conceitos básicos em Teleprocessamento	57
2.9. Microcomputadores	58
2.10. Desenvolvimento de sistemas	60
3. INFORMÁTICA APLICADA	65
3.1. Automação dos serviços de uma biblioteca	65
3.1.1. Aquisição	67

3.1.2.	Emissão de catálogos	68
3.1.3.	Circulação	70
3.2.	A documentação automatizada	71
3.2.1.	Indexação automatizada	71
3.2.2.	Recuperação automática de informações	74
4.	ANEXOS	85
4.1.	Anexo 1 — O ILS	85
4.2.	Anexo 2 — O DOBIS	92
4.3.	Anexo 3 — O Videotexto	95
4.4.	Anexo 4 — O ENDS	97
4.5.	Anexo 5 — O STAIRS	106
4.6.	Anexo 6 — Informática e Cultura: um exemplo de pioneirismo	111
4.7.	Anexo 7 — Glossário de termos, siglas e acrônimos ..	116
4.8.	Anexo 8 — Miniléxico inglês-português	120
4.9.	Anexo 9 — Uma incursão no amanhã (a guisa de conclusão)	124

1

INTRODUÇÃO

"Um bibliotecário não precisa transformar-se num programador ou num analista de sistemas para desempenhar bem suas funções."

Este trabalho pretende juntar noções teóricas, colocações práticas e resultados de experiências pessoais, visando formar um conjunto de informações que possa ser utilizado como ponto de partida para estudos mais profundos e especializados.

Vários anos lecionando informática (sem entrar no mérito da propriedade deste termo, de várias acepções e usos) em escolas de biblioteconomia, organizando e ministrando cursos de reciclagem, bem como exercendo uma atividade profissional voltada ao manuseio de informações, via computador, nos permitem avaliar as dificuldades encontradas pelos alunos e pelos profissionais de biblioteconomia, quando se deparam com o computador pela primeira vez.

Mesmo após um contato inicial, via de regra pelo manuseio de algum microcomputador e de algum programa pré-fabricado, as dificuldades continuam e no momento em que o bibliotecário é obrigado, muitas vezes à revelia, a usar o computador, quer seja na automação de tarefas biblioteconômicas, quer seja na utilização de bases de dados, a indecisão, a perplexidade e um certo estado de impotência se instalam neste profissional, o qual começa, freneticamente, a devorar informações sobre o assunto e a recorrer aos mais diversos especialistas em informática, buscando a fórmula mágica.

Mas não há fórmulas mágicas. O que há é um estudo consciente e sistemático sobre o assunto, que deve começar antes da ne-

cessidade se configurar, acompanhado de um espírito crítico e aberto, aliado a muito bom senso e reflexão.

A adoção de “ferramentas eletrônicas” no desempenho das atividades profissionais requer mais mudanças de comportamento do que aprendizado tecnológico.

É evidente que se este último existir, tanto melhor. Entretanto, para um bibliotecário desempenhar bem suas funções, fazendo uso dos meios mais modernos ao seu alcance, não é preciso que ele se transforme em profissional de outras áreas.

Em outras palavras: um bibliotecário não precisa se tornar um programador nem um analista de sistemas, e muito menos um engenheiro eletrônico, para desempenhar suas funções a contento.

Os conhecimentos de planilhas eletrônicas, de editores de textos e dos D. Bases da vida não são nocivos, pelo contrário. Mas, se o bibliotecário achar que somente com uma planilha eletrônica e um D. Base III Plus, ele vai resolver seus problemas profissionais, é provável que ele tenha várias decepções e seus usuários mais ainda.

Assim sendo, não vamos esperar milagres deste livro (nem de nenhum outro). Vamos lê-lo com os olhos e o espírito abertos, procurando não se fixar no detalhe, mas tentando captar-lhe a mensagem geral.

Como diríamos em “computês”: não nos fixemos no BIT, processemos o arquivo.

2 INFORMÁTICA TEÓRICA

"Subestimar o computador é um erro. Superestimá-lo pode ser fatal."

2.1. O COMPUTADOR

2.1.1. Breve histórico

Os computadores, como não podia ser de outra forma, não nasceram assim como os conhecemos atualmente. Eles foram evoluindo no decorrer dos tempos, partindo de concepções primitivas e chegando ao estágio atual que, por sua vez, está longe de ser o ponto final desta evolução assombrosa. Nenhum produto tecnológico evoluiu tanto e em tão pouco tempo como os computadores eletrônicos. Além das mudanças quantitativas houve significativas mudanças qualitativas que deram origem às chamadas gerações de computadores. Neste breve histórico abordaremos superficialmente alguns fatos ou descobertas entre o ábaco e o primeiro computador a merecer este nome, o ENIAC. As gerações de computadores serão tratadas adiante.

Após um período inicial onde o homem fazia uso dos próprios dedos para efetuar cálculos elementares, auxiliares mecânicos foram sendo criados e desenvolvidos para fazer frente às necessidades que surgiam. Um dos primeiros auxiliares mecânicos foi o ábaco, que

alguns autores datam do século III a.C., aproximadamente, atribuindo esta invenção ora a chineses ora a babilônios ou outras culturas da época. Ainda hoje é possível ver tipos aprimorados destes instrumentos sob forma de esferas coloridas enfiadas em arames.

Depois de um longo período sem maiores novidades, os progressos da matemática e da astronomia fizeram com que novos métodos de cálculo fossem buscados.

Por volta de 1614, o matemático escocês John Napier concebeu uma forma de simplificar as operações de multiplicação e de divisão, difíceis de serem efetuadas com o ábaco, inventando os logaritmos. Com isso, as multiplicações tornaram-se adições de logaritmos e as divisões subtrações de logaritmos. Esta invenção proporcionou o aparecimento da régua de cálculo.

Em 1644, aproximadamente, o matemático e filósofo francês Blaise Pascal inventou a primeira máquina para efetuar somas e subtrações, usando o princípio de rodas dentadas acopladas de tal forma que qualquer uma delas, ao superar a contagem de nove, transmitia à roda seguinte, de ordem mais alta, o “vai um” correspondente do sistema decimal. Analogamente, o processo funcionava na subtração com a transmissão no sentido das rodas de ordem mais baixa.

Por volta de 1833 o matemático inglês Charles Babbage projetou os planos para a primeira calculadora automática do mundo, controlada por programa.

Influenciado pela invenção de Joseph Marie Jacquard, que em 1801 construíra um tear controlado por cartões perfurados, Babbage desenvolveu a Máquina Analítica que pode ser considerada a precursora dos modernos computadores digitais. Como estes, a Máquina Analítica possuía os elementos: Entrada, Memória, Processamento e Saída, sendo sua operação governada por cartões perfurados. A máquina de Babbage dispensava qualquer intervenção humana, a menos que viesse a necessitar de informações adicionais, podendo, inclusive, alterar o andamento dos cálculos de acordo com os resultados intermediários obtidos. Entretanto, esta máquina não chegou a ser concretizada por limitações da tecnologia da época e Babbage faleceu em 1871 sem poder completar seu trabalho. Contudo, o trabalho não foi perdido. Progressos significativos ocorreram após a morte de Babbage e o mais importante foi obtido por Herman Hollerith.

Por ocasião do censo norte-americano de 1880, 500 funcionários, sob as ordens do dr. Hollerith, trabalharam durante 7 anos para apurar os resultados contidos em 55 milhões de questionários. Diante disto, Hollerith criou um cartão padronizado para cada pessoa. Este cartão, contendo os dados a serem levantados em posições fixas e predeterminadas, seria perfurado nos lugares correspondentes e esta perfuração representaria a resposta SIM, em contraposição à ausência de perfuração que indicaria a resposta NÃO. Feito isto, os cartões eram lidos por uma máquina, também de invenção de Hollerith, cujo princípio de funcionamento era relativamente fácil. O cartão passava entre dois cilindros, dos quais o superior tinha uma pequena escova metálica que varria o cartão em toda sua largura. Nos lugares do cartão onde houvesse uma perfuração, os fios metálicos da escova entravam em contato com o cilindro metálico inferior e um mecanismo contador era acionado. Com esta invenção o tempo de apuração do censo de 1890 foi reduzido drasticamente (cerca de um terço do anterior), embora a população tivesse aumentado em mais de 10 milhões. Mais tarde, Hollerith fundou uma empresa para fabricar e distribuir suas máquinas, absorvida em 1911 pela Computing Tabulating Recording Company, cujo nome foi alterado em 1924 para International Business Machine Corporation, a conhecida IBM.

A iminência da eclosão da Segunda Grande Guerra acelerou as pesquisas nesta área e já em 1941 o alemão Conrad Zuse constrói a primeira calculadora automática do mundo, a Z3, controlada por programa, que funcionou razoavelmente.

Em 1943, Howard Aiken desenvolveu para o governo americano o Mark I, que tinha um funcionamento similar ao da Z3.

O primeiro computador eletrônico, assim chamado devido à utilização de válvulas eletrônicas em lugar de relés eletromagnéticos, foi construído, em 1945, na Universidade de Pensilvânia, pelos professores John Eckert e John Mauchly e denominado ENIAC (Electronic Numeric Integrator And Calculator). O ENIAC pesava 30 toneladas, ocupava 1.400m² de superfície e utilizava milhares de km de fios. Tinha 18.000 válvulas eletrônicas e trabalhava com o sistema decimal. Sua velocidade de cálculo era muito maior que a das máquinas anteriores, mas tinha, como inconveniente, a grande taxa de manutenção e uma programação pouco flexível, efetuada por nu-

merasas ligações em painéis diferentes para cada problema a ser solucionado.

Os computadores com a filosofia de programa com armazenamento interno das instruções surgiram em 1949, graças ao trabalho do dr. John Von Neumann, em Cambridge.

2.1.2. Gerações

Como vimos no item anterior o ENIAC pode ser considerado o primeiro computador eletrônico que merece este nome. Do ENIAC, que foi usado entre 1945 e 1958 aproximadamente, até os modernos computadores de hoje, muitas coisas mudaram. A tecnologia dos componentes eletrônicos avançou vertiginosamente e isso possibilitou a construção de computadores cada vez mais rápidos, mais possantes, de menor tamanho e de custos relativamente mais baixos.

A seguir vamos apresentar esquematicamente a trajetória destes computadores, dividindo-os em 5 intervalos de tempo que representam as mudanças qualitativas. A cada intervalo de tempo que caracteriza as grandes mudanças é dado o nome, bem apropriado, de geração de computadores e as datas-limites de cada intervalo são aproximadas.

A primeira geração pode ser situada entre 1951 e 1958 e suas principais características são:

- circuitos com válvulas e quilômetros de fios;
- velocidade de processamento medida em milissegundos;
- muito caros;
- muito grandes;
- necessidade de manutenção excessiva;
- programação por painéis externos.

A segunda geração vai de 1958 a 1965 e se caracteriza por:

- circuitos impressos e transistorizados;
- velocidade de processamento medida em microssegundos;
- relativamente baratos;
- dimensões mais reduzidas;
- menor necessidade de manutenção;
- programação usando linguagens simbólicas.

A terceira geração que vai de 1965 a 1970 apresenta:

- circuitos monolíticos integrados;
- programação usando linguagens mais fáceis e poderosas;
- aparece o conceito de multiprogramação, pelo qual vários programas podem ser executados concorrentemente, ficando o controle da execução a cargo do próprio computador;
- velocidade de processamento medida em nanossegundos.

A quarta geração têm como características principais:

- circuitos monolíticos integrados em larga escala;
- velocidade de processamento medida em picossegundos;
- programação usando linguagens conversacionais;
- aparecimento dos microcomputadores;
- construção de redes de microcomputadores distribuídos.

Esta geração pode ser situada entre 1970 e os dias atuais, pelo menos em termos comerciais, muito embora desde 1985 a quinta geração já não seja apenas um projeto, pelo menos no Japão e nos Estados Unidos.

Quinta geração — o futuro (presente):

- inteligências artificiais;
- robótica aprimorada;
- máquinas mais voltadas ao raciocínio do que ao processamento;
- computadores orgânicos com microprocessadores biológicos em lugar de pastilhas de silício;
- pela engenharia genética serão criadas moléculas de proteínas com capacidade de armazenar e processar informações.

Nesta breve incursão no campo da evolução dos computadores ficou evidenciada a verdadeira revolução que determinou o fim de uma geração e o começo da seguinte.

A tendência atual é fazer evoluir paralelamente a máquina e o instrumental que permite utilizá-la, como, por exemplo, as linguagens e os sistemas operativos. Isto porque, a rigor, a máquina em si tem um potencial muito maior do que aquele efetivamente usado.

Daí o grande esforço em desenvolver linguagens realmente conversacionais, sistemas operativos possantes e flexíveis, “pacotes”

prontos para a solução de problemas específicos e para cujo uso não seja necessário ser iniciado, e assim por diante.

Esta parece ser a tendência e ela só pode trazer benefícios para os usuários em geral, dentre os quais os bibliotecários se constituem num segmento cada vez mais significativo.

2.1.3. Tipos

Os computadores podem ser classificados segundo alguns critérios.

Se o critério for capacidade de processamento, temos computadores de pequeno, médio e grande porte e ainda mini e micro-computadores.

Se o critério for características de operação, temos os computadores analógicos e digitais. Os computadores analógicos representam as quantidades por meio de grandezas físicas, ao passo que os digitais representam estas quantidades por meio de símbolos (letras, números e caracteres especiais). Para termos uma idéia clara da diferença entre uma representação analógica e uma digital, basta comparar a forma de representar a hora entre um relógio analógico e um relógio digital. Podemos dizer que o computador analógico mede e o digital conta.

Vistos sob o prisma da aplicação, temos computadores para fins científicos e computadores para fins comerciais. As aplicações científicas se caracterizam por pequenas quantidades de dados e cálculos enormes e complexos. Isto quer dizer que um computador voltado para este tipo de aplicação deve ter a parte de processamento altamente desenvolvida, podendo ser a capacidade de entrada e saída relativamente pequena. Já um computador voltado para aplicações comerciais deve ter grande capacidade de entrada e saída, podendo efetuar processamentos relativamente simples. Isto porque as aplicações comerciais têm esta característica.

2.1.4. Componentes

Como foi dito anteriormente, o computador é um sistema e, como tal, é constituído por um conjunto de elementos que lhe permitem alcançar seus objetivos.

Os elementos que constituem o grande sistema podem ser divididos, num primeiro momento, em duas grandes categorias.

Os elementos físicos, que são as próprias máquinas, cada uma com uma finalidade específica, e os elementos que podemos chamar de lógicos, os quais compreendem essencialmente os programas, que fazem a máquina funcionar e conseguir os resultados desejados.

O primeiro grupo, o das máquinas, é denominado genericamente de **HARDWARE**, cabendo ao segundo a denominação de **SOFTWARE**.

Assim sendo, uma conceituação possível de computador pode ser enunciada como segue: "computador é um conjunto de máquinas, ou hardware, inter-relacionadas, com a finalidade de processar dados e/ou informações, utilizando programas, ou software, que definem o processamento a ser efetuado".

Neste item serão apresentados os componentes físicos do computador, o hardware.

O software será abordado mais adiante ao falarmos sobre programação e linguagens, desenvolvimento de sistemas e outros assuntos não ligados à maquinaria.

O hardware de um sistema de computador pode ser separado em três blocos fundamentais, segundo a finalidade:

- unidades de entrada de dados;
- unidades de processamento de dados e
- unidades de saída de dados.

No bloco das unidades de entrada estão agrupadas todas as máquinas que, de uma forma ou de outra, permitem aos dados a serem processados entrarem no sistema.

As unidades de processamento têm a finalidade de processar estes dados de acordo com as especificações dos programas.

As unidades de saída por sua vez permitem devolver ao meio externo os resultados do processamento efetuado.

As unidades de entrada e de saída são conhecidas também pelo nome de **periféricos**, por estarem, de certo modo, ao redor da unidade central.

A arquitetura de cada periférico depende do tipo de veículo de dados a ele destinado. Uma impressora, por exemplo, tem sua arquitetura voltada para a impressão de caracteres em papel, que é o veículo dos dados.

A quantidade de periféricos que podem ser conectados à unidade central depende das necessidades de quem for usar o computador, embora existam limites estabelecidos pela própria tecnologia de cada modelo. Porém, dentro destes limites, bastante amplos, é possível acomodar ou retirar periféricos segundo as necessidades.

Graças a esta característica, os sistemas de computador são sistemas modulares.

Tendo em vista a finalidade e o alcance desta publicação, é de pouca ou nenhuma utilidade apresentar detalhes técnicos sobre as máquinas responsáveis pela entrada e saída de dados no e do computador.

Consideramos mais oportuno conhecer os veículos sobre os quais os dados circulam, tendo sempre em mente que para cada um deles existe uma máquina apropriada da qual se retiram (se for uma unidade de entrada) ou na qual se registram (se for uma unidade de saída) dados e/ou informações.

Assim sendo conheceremos estas unidades através dos veículos que delas se utilizam para introduzir ou retirar dados do computador.

Terminando a apresentação dos veículos, descreveremos a unidade central de processamento, encerrando assim esta parte relativa ao hardware de um sistema de computador.

2.1.4.1. Veículos de dados e/ou informações

— Cartão perfurado

Foi o primeiro veículo de dados utilizado pelos computadores e durante muitos anos foi empregado em grande escala.

Atualmente seu uso é muitíssimo reduzido e a tendência é desaparecer de vez.

Trata-se de um retângulo de cartolina no qual, através de máquinas apropriadas, são efetuadas perfurações que obedecem um código específico. Nesse código, as perfurações, dependendo de sua localização e quantidade, representam letras, algarismos, sinais especiais, etc.

Assim, os dados a serem processados são transcritos de seus documentos originais para esses cartões, os quais são lidos por um periférico específico do computador, chamado unidade leitora de

cartões, que decodifica as perfurações e envia os dados à unidade central de processamento para serem processados.

O próprio computador tem condições de devolver resultados do processamento, perfurando-os em cartões, através da unidade perfuradora de cartões, os quais poderão voltar a ser lidos pelo computador se e quando necessário.

Assim, é fácil verificar que o cartão perfurado é, ao mesmo tempo, um veículo de entrada e de saída, dependendo da necessidade do momento.

— *Fita de papel perfurada*

Também em desuso, tem uma filosofia de funcionamento análoga à do cartão perfurado.

Aqui, as perfurações são dispostas ao longo de uma tira de papel que é lida por uma unidade especial, podendo também ser perfurada por outra unidade do computador.

Isto torna a fita de papel também um veículo de entrada e saída.

A fita de papel perfurada foi largamente empregada em bibliotecas que efetuavam o desdobramento de fichas catalográficas usando um equipamento denominado Flexowriter, que permitia, por meio de um equipamento acoplado denominado Selectadata, reproduzir n fichas a partir de uma ficha-matriz.

A fita de papel, perfurada simultaneamente à datilografia da ficha-matriz, era lida várias vezes e a cada vez era gerada uma ficha de acordo com as especificações do programa, por sua vez perfurado em outra fita de papel e interpretado pela Selectadata.

Algumas bibliotecas usaram essas fitas perfuradas, contendo seu acervo, ou parte dele, por ocasião da automação de seus procedimentos, evitando, desta forma, o enorme trabalho de digitação e conferência dos dados.

Para isso, evidentemente, o computador devia possuir o periférico apropriado para ler fitas de papel perfurado.

— *Fitas magnéticas*

Ao contrário das fitas de papel, as fitas magnéticas são amplamente utilizadas nos modernos computadores.

Esta é a grande vantagem dos discos sobre as fitas, que justifica seu uso apesar de seu custo mais elevado.

Existem aplicações em que a rapidez na obtenção das informações é primordial. Neste caso, o uso de discos magnéticos é imperativo.

Os discos magnéticos também podem ser regravados.

— Discos óticos

Conhecidos também como discos lasers, permitem armazenar enormes quantidades de dados.

Recentemente comercializados, os discos óticos usam raios laser para a gravação dos dados. Esta tecnologia de gravação permite uma densidade de gravação muitas vezes maior que aquela obtida nos discos magnéticos convencionais.

O inconveniente desse tipo de discos está relacionado com as dificuldades de regravação, sendo que, até o momento, eles são recomendados para aplicações onde apenas a leitura dos dados seja requerida.

O impacto desses discos na biblioteca e no centro de documentação é enorme. A capacidade de armazenamento é tão grande que bancos de dados inteiros podem ser estocados neles. Isto causa uma revolução na medida em que diminui ou mesmo evita o acesso a certas bases de dados, nacionais ou internacionais, via teleprocessamento (com todos os inconvenientes inerentes a este tipo de acesso), substituindo-o pela aquisição de coleções completas, gravadas nestes CD-ROM (*compact disk-read only memory*), como também são conhecidos. As vantagens são óbvias: o processamento destas informações pode ser efetuado no próprio local, da forma mais conveniente e utilizando o *software* mais adequado ou disponível; não há o problema das linhas que “caem”, deixando o usuário a “ver navios”, ou mesmo o problema de “tempos de resposta” que mais se assemelham a “prazos de entrega”.

Vários Sistemas Informativos já trabalham neste esquema de divulgação das informações.

— Documentos originais

Existem aplicações nas quais é conveniente e possível processar dados que já constem do documento original, evitando o processo de

Esta é a grande vantagem dos discos sobre as fitas, que justifica seu uso apesar de seu custo mais elevado.

Existem aplicações em que a rapidez na obtenção das informações é primordial. Neste caso, o uso de discos magnéticos é imperativo.

● Os discos magnéticos também podem ser regravados.

— Discos óticos

Conhecidos também como discos lasers, permitem armazenar enormes quantidades de dados.

Recentemente comercializados, os discos óticos usam raios laser para a gravação dos dados. Esta tecnologia de gravação permite uma densidade de gravação muitas vezes maior que aquela obtida nos discos magnéticos convencionais.

O inconveniente desse tipo de discos está relacionado com as dificuldades de regravação, sendo que, até o momento, eles são recomendados para aplicações onde apenas a leitura dos dados seja requerida.

O impacto desses discos na biblioteca e no centro de documentação é enorme. A capacidade de armazenamento é tão grande que bancos de dados inteiros podem ser estocados neles. Isto causa uma revolução na medida em que diminui ou mesmo evita o acesso a certas bases de dados, nacionais ou internacionais, via teleprocessamento (com todos os inconvenientes inerentes a este tipo de acesso), substituindo-o pela aquisição de coleções completas, gravadas nestes CD-ROM (*compact disk-read only memory*), como também são conhecidos. As vantagens são óbvias: o processamento destas informações pode ser efetuado no próprio local, da forma mais conveniente e utilizando o *software* mais adequado ou disponível; não há o problema das linhas que “caem”, deixando o usuário a “ver navios”, ou mesmo o problema de “tempos de resposta” que mais se assemelham a “prazos de entrega”.

Vários Sistemas Informativos já trabalham neste esquema de divulgação das informações.

— Documentos originais

Existem aplicações nas quais é conveniente e possível processar dados que já constem do documento original, evitando o processo de

transcrição. É o caso de cheques, cartões para uso de bancos 24 horas, cartões da loto, códigos de barras em produtos de supermercados, etc.

Neste caso, o documento é lido por máquinas apropriadas e os dados existentes, em locais prefixados, são enviados diretamente ao complexo central para processamento.

Os dados inteligíveis ao computador, existentes nesta categoria de documentos, podem ser de origem magnética ou ótica, recebendo a denominação de caracteres magnéticos, ou MCR, e caracteres óticos, ou OCR, respectivamente.

— *Digitalizadores de imagens*

Permitem converter imagens (fotos, pinturas, mapas, etc.) numa forma inteligível ao computador.

Um dispositivo apropriado varre a imagem transformando-a em informação digital e gravando-a em disco para uso posterior.

Através de programas especiais, essas imagens, gravadas em forma digital, podem ser exibidas em telas e/ou impressas.

— *Mesas para gráficos*

Dispositivos manuais conectados a computadores que permitem, por meio de uma caneta apropriada, efetuar desenhos, obtendo-se o resultado diretamente na tela.

Certos comandos ou instruções fornecidos ao computador permitem inverter o desenho, aumentá-lo ou diminuí-lo e outras facilidades que tornam este processo muito importante para laboratório de desenho, indústrias automobilísticas, etc.

— *Teclados*

Permitem introduzir dados no computador, digitando-os.

Trata-se de uma forma de entrada de dados largamente utilizada, quer seja em microcomputadores ou em terminais conectados a computadores de grande porte.

— *Impressos*

Sob esta designação genérica estão reunidas todas as saídas

impressas em formulários, sejam do tipo que forem: relatórios, cartões, extratos, avisos, etc.

Para obter este tipo de saída existem equipamentos que vão desde modestas impressoras de mesa, relativamente lentas e pobres em recursos, até possantes impressoras a laser com velocidade de impressão espantosa e sofisticados recursos adicionais.

— Vídeos

Semelhante à tela de um televisor, este dispositivo apresenta na tela os resultados do processamento efetuado, bem como os dados digitados.

Este vídeo acoplado a um teclado permite uma interação perfeita entre o homem e o computador. Considerando a enorme velocidade de processamento, o tempo entre o envio dos dados a serem processados, via teclado, e o recebimento das respostas, via vídeo, é cada vez menor, estabelecendo-se um verdadeiro diálogo entre o homem e a máquina.

— Memórias auxiliares

A razão de tratarmos das memórias auxiliares neste item, destinado aos veículos de dados, se deve ao fato de que uma memória auxiliar pode tornar-se, a qualquer momento, um veículo de entrada e/ou saída para, em seguida, voltar a ser uma memória auxiliar.

Exemplificando: uma fita magnética que está sendo gravada por um periférico específico do computador é, nesse momento, um veículo de saída. Terminada a gravação e guardada a fita magnética numa estante, ela se torna uma memória auxiliar que armazena dados pelo tempo necessário. Ao retirar essa fita da estante, para colocá-la numa leitora de fitas magnéticas, com o intuito de processar os dados nela gravados, ela passa a ser um veículo de entrada.

Assim, em princípio, toda memória auxiliar é, em determinadas circunstâncias, um veículo de entrada e/ou saída e vice-versa.

2.1.4.2. Unidade Central de Processamento

A unidade genericamente denominada Unidade Central de Processamento, também conhecida pelas siglas UCP ou CPU (de

Central Processing Unity)), na verdade é constituída de três módulos, que são:

— Unidade Aritmética e Lógica (UAL), que contém circuitos eletrônicos especiais que lhe permitem efetuar operações aritméticas (somas, subtrações, divisões e multiplicações) e operações lógicas (comparações, exclusões, ordenações, etc.).

— Unidade de Controle (UC), responsável pelo controle de todo o sistema de computador.

— Unidade de Memória Central ou Principal (UMP), à qual cabe armazenar programas e dados a serem processados. Uma vez terminado o processamento, essa memória é liberada para receber outros dados. A memória principal não é depósito de dados. Esse papel é reservado às memórias auxiliares.

A Unidade Central de Processamento pode ser considerada o "centro nervoso" de um computador. Nenhum processamento, por mais elementar que seja, pode ser efetuado sem a participação desta.

O exemplo a seguir procura mostrar o funcionamento de uma UCP e seu relacionamento com os periféricos.

Suponhamos ter de somar entre si um conjunto de números gravados em fita magnética e imprimir o resultado. Os passos para resolver este problema são:

1º) O sistema é acionado pelo operador.

2º) O programa que deve resolver o problema é gravado na UMP, na parte reservada aos programas.

3º) A UC vai buscar a primeira instrução do programa.

4º) Essa instrução manda ler a fita magnética e descarregar os dados desejados (no caso os números a serem somados) na UMP, na parte reservada aos dados.

5º) Pela segunda instrução lida, a UC fica sabendo que deve introduzir na UAL o primeiro dos números a serem somados.

6º) A próxima instrução transfere o segundo número a ser somado da UMP para a UAL, onde é somado ao primeiro que lá se encontrava. E assim sucessivamente, até terminarem os números, quando então o resultado é guardado na UMP até saber o que fazer com ele.

7º) Isso é determinado pela próxima instrução que manda imprimir o resultado em papel. Assim, a UC transfere para a impressora o resultado que se encontrava na UMP.

8º) A impressora vai imprimir esse resultado obedecendo as especificações de impressão constantes no programa.

Pelo exemplo acima é possível ter uma idéia do funcionamento da Unidade Central de Processamento, bem como perceber que todo e qualquer dado a ser processado, esteja no veículo que estiver, deve ser transferido primeiramente para a Memória Principal para em seguida ser processado. Da mesma forma, os resultados devem ser gravados na Memória Principal para depois serem destinados aos periféricos especificados.

2.2. SISTEMAS DE NUMERAÇÃO

Uma das perguntas mais freqüentes e angustiantes que nos têm sido feitas pelos estudantes de biblioteconomia diz respeito à necessidade de conhecimentos matemáticos para o aprendizado da informática. Com essa pergunta o aluno quer saber até que ponto ele precisa conhecer matemática para acompanhar as aulas e até que ponto serão dados ensinamentos matemáticos no decorrer do curso. Essa preocupação deve estar presente também em boa parte dos eventuais leitores deste livro. A estes respondemos o mesmo que respondemos àqueles, ou seja: a única matemática de que vamos precisar é a das quatro operações aritméticas, quando falarmos de sistemas de numeração, nada mais. Mesmo porque, se fosse preciso mais, seria preciso também outro professor. Sendo assim, chegou o momento de lembrarmos a tabuada e usá-la para entendermos sistemas de numeração.

Quando o homem descobriu as noções de quantidade no seu dia-a-dia, as transações envolviam pequenos números capazes de serem representados pelos dez dedos das mãos. Cada um destes números recebeu um nome e, posteriormente, um símbolo que o representava.

Com o desenvolvimento da civilização e o aumento das transações, os dez números existentes se tornaram insuficientes e, por outro lado, era inviável criar um nome e uma representação particular para cada número que viesse a ser criado.

Assim, foram elaborados sistemas que permitem formar qualquer número desejado a partir de uns poucos símbolos básicos.

Para se ter uma idéia de como esse reaproveitamento é eficiente, basta dizer que com apenas 10 símbolos podemos criar infinitos números, rigorosamente falando.

Assim surgiram os sistemas de numeração, dentre os quais o mais conhecido e utilizado é o sistema decimal.

Podemos conceituar sistema de numeração como sendo um agrupamento ordenado de números no qual uma determinada quantidade de unidades de uma ordem origina uma unidade de ordem imediatamente superior.

O número de unidades de uma ordem que origina uma unidade de ordem imediatamente superior recebe o nome de base do sistema de numeração.

Um sistema de numeração é regido por duas regras gerais:

1ª) Num sistema qualquer de numeração de base N , são necessárias N unidades de uma determinada ordem para originar uma unidade de ordem imediatamente superior.

2ª) Num sistema qualquer de numeração de base N , um algarismo colocado imediatamente à direita de um outro, vale N vezes menos do que este último. (Ao dizermos à direita de um outro queremos dizer à nossa direita. Ex: considerando dois algarismos XY , o Y está à direita do X .)

Cada símbolo dentro de um número tem dois valores: o valor intrínseco, por ser o que ele é independentemente do contexto, e o valor posicional, devido à posição que ele ocupa dentro do número.

O valor posicional é obtido multiplicando-se o valor intrínseco pelo peso da posição da casa ocupada pelo símbolo, sendo que os pesos das sucessivas casas, da direita para a esquerda, são as sucessivas potências da base: $N^0, N^1, N^2, N^3, \dots, N^n$.

Dentre os sistemas de numeração existentes, três são mais importantes para o nosso estudo:

Sistema decimal	— de base 10,
Sistema binário	— de base 2,
Sistema hexadecimal	— de base 16.

Sistema decimal

É o mais conhecido e mais usado sistema de numeração, de modo que será o primeiro a ser apresentado.

No sistema decimal, dez unidades de uma certa ordem constituem uma unidade de ordem imediatamente superior.

A base do sistema decimal é 10.

Neste sistema qualquer número pode ser escrito com os símbolos: 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9, menores que a base.

No sistema decimal, o número quinhentos e cinquenta e cinco é representado por 555, onde cada símbolo 5, embora possuindo o mesmo valor intrínseco, difere dos demais pelo seu valor posicional, ou seja, pela posição que ele ocupa dentro do número todo. Por se tratar de um sistema de base 10, o símbolo 5 que se encontra no meio é dez vezes menor do que o 5 à sua esquerda e dez vezes maior que o 5 à sua direita. Isto quer dizer que o número 555 pode ser decomposto como segue: 5 (centena) + 5 (dezena) + 5 (unidade), que é o mesmo de $5 \cdot 10^2 + 5 \cdot 10^1 + 5 \cdot 10^0 \longrightarrow 5 \cdot 100 + 5 \cdot 10 + 5 \cdot 1 \longrightarrow 500 + 50 + 5 = 555$.

Sistema binário

É um sistema de numeração que se difundiu muito após o advento dos computadores.

Neste sistema, duas unidades de uma certa ordem constituem uma unidade de ordem imediatamente superior.

A base do sistema é 2.

Qualquer número deste sistema pode ser representado com os símbolos 0 e 1, menores que a base.

Analogamente ao sistema decimal, no número 111, o 1 do meio é duas vezes menor que o 1 à sua esquerda e duas vezes maior que o 1 à sua direita. O número 111 pode ser decomposto da seguinte forma: $1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 \longrightarrow 1 \cdot 4 + 1 \cdot 2 + 1 \cdot 1 \longrightarrow 4 + 2 + 1 = 7$.

Sistema hexadecimal

Este sistema também deve sua difusão aos computadores, que o empregam largamente.

No sistema hexadecimal, 16 unidades de uma certa ordem constituem uma unidade de ordem imediatamente superior.

A base deste sistema é 16.

Qualquer número deste sistema pode ser representado com os símbolos: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E e F, menores que a

base (onde A, B, C, D, E e F correspondem a 10, 11, 12, 13, 14 e 15, no sistema decimal).

Analogamente aos dois sistemas anteriores, no número hexadecimal AAA, o A do meio é 16 vezes menor que o A à sua esquerda e 16 vezes maior que o A à sua direita. Assim ele pode ser decomposto desta forma: $A.16^2 + A.16^1 + A.16^0 \longrightarrow A.256 + A.16 + A.1$ (lembrando que $A = 10$) $\longrightarrow 10.256 + 10.16 + 10.1 \longrightarrow 2.560 + 160 + 10 = 2.730$.

2.3. REPRESENTAÇÃO DAS INFORMAÇÕES NO COMPUTADOR

Como já sabemos, o sistema de computador recebe dados do meio externo, através de um ou mais periféricos de entrada, processa-os e devolve os resultados ao meio externo por meio de um ou mais periféricos de saída. Entretanto, para manipular dados é preciso que os mesmos possam ser representados.

O ser humano consegue manipular números, por exemplo, porque estes podem ser representados pelos símbolos: 0, 1, 2, 3, 4, 5, 6, 7, 8 e 9. Isto permite ao ser humano efetuar contas, comparações e outros processamentos quaisquer.

As unidades do computador encarregadas de efetuar os processamentos são constituídas por circuitos eletrônicos. De que modo podemos representar as informações num meio eletrônico?

A representação é feita por meio de grandezas físicas passíveis de serem reconhecidas pela máquina. Assim sendo, o registro das informações é feito por meio de tensões elétricas e/ou magnetismo, que permitem maior rapidez de leitura e de gravação.

O computador trabalha de modo binário, isto é, ele somente pode representar símbolos (letras, números e caracteres especiais) por meio da combinação de dois estados possíveis.

A passagem de corrente por um fio é um estado que se opõe à não passagem de corrente.

A magnetização de um metal é um estado em oposição à não magnetização.

É assim que o computador trabalha.

Mas como é possível representar todos os números, letras e demais caracteres utilizados na comunicação humana? Isto implica

a necessidade de um código que permita combinar estes dois estados para obter os símbolos desejados.

Para exemplificar suponhamos uma torneira, a qual pode estar aberta ou fechada.

Se houvesse apenas dois símbolos, "a" e "b", para representar, seria fácil. Bastaria convencionar que torneira aberta representaria "a" e torneira fechada representaria "b". Mas não existem apenas as letras "a" e "b". E as demais?

A matemática nos ensina que uma única torneira, podendo assumir dois estados distintos (aberta ou fechada), nos dá 2^1 possibilidades de representação: isto é, 2 representações.

Se dispuséssemos de duas torneiras, o número de possibilidades aumentaria para 2^2 ; isto é, 4 representações.

Para facilitar o raciocínio, vamos representar a torneira aberta por "1" e a torneira fechada por "0".

As combinações obtidas com duas torneiras seriam: 00, 01, 10 e 11 (ambas fechadas, a primeira fechada e a segunda aberta e assim por diante), que poderiam representar as letras "a", "b", "c" e "d", respectivamente.

Mesmo assim, ainda faltam muitas torneiras para as nossas necessidades de representação:

3 torneiras permitiriam $2^3 = 8$ possibilidades,
4 torneiras permitiriam $2^4 = 16$ possibilidades,
e assim sucessivamente.

Isto mostra como é possível representar um conjunto grande de símbolos a partir de dois apenas, desde que combinados dentro de um código que estabeleça as relações entre cada combinação e aquilo que ela deve representar.

Este código, como todos, é regido por leis que atribuem à mensagem um mesmo valor, tanto para o emissor como para o receptor da mesma.

No computador não há torneiras. A memória principal de um computador é constituída por pequenas argolas metálicas magnetizáveis denominadas ferrites ou núcleos magnéticos.

Cada argola pode assumir dois estados possíveis: magnetizada ou não magnetizada. Por isso estas argolas recebem a designação de dígitos binários, mais conhecidos por BITS, do inglês BInary digiT.

Quanto maior for o número de argolas, maior será a capacidade de armazenamento da memória principal e, conseqüentemente, de processamento.

Assim, como vimos no exemplo das torneiras, as memórias principais usam um conjunto de ferrites ou BITS para representar os símbolos da comunicação humana.

É desta forma, portanto, que as informações são representadas no computador, obedecendo determinados códigos, conforme apresentado a seguir.

2.4. PRINCIPAIS CÓDIGOS UTILIZADOS PELOS COMPUTADORES

O componente básico da informação armazenada na memória de um computador é o dígito binário, designado pelo nome técnico de BIT.

Existem várias maneiras de se agrupar os bits para obter as informações. Cada maneira constitui um código diferente.

Para uso em computadores temos, basicamente, quatro códigos distintos:

- Binário n bits
- Binário-decimal (BCD) 6 bits
- Binário-decimal estendido (EBCDIC) 8 bits
- Americano padronizado (ASCII) 8 bits

Código binário

Este código nada mais é do que o próprio sistema de numeração de base 2, já visto:

$2^0 = 1$	$2^4 = 16$	$2^8 = 256$
$2^1 = 2$	$2^5 = 32$	$2^9 = 512$
$2^2 = 4$	$2^6 = 64$	$2^{10} = 1.024$
$2^3 = 8$	$2^7 = 128$	• = •
		• = •
		• = •
		$2^n = n$

Código BCD

É uma derivação do código binário que oferece maior facilidade de manuseio.

O princípio do código BCD é muito simples; escreve-se o número sob forma decimal e em seguida codifica-se cada um dos algarismos em código binário puro.

Vejamos, por exemplo, como seria escrito o número 7.940 em BCD:

7	9	4	0
0111	1001	0100	0000

No exemplo cada algarismo foi representado por 4 bits porque, como vimos, 4 bits são suficientes para representar os 10 algarismos decimais (e ainda sobra).

Entretanto, nos computadores que trabalham com este código, cada posição de memória é constituída por 6 bits porque o computador deve poder representar também letras e sinais especiais.

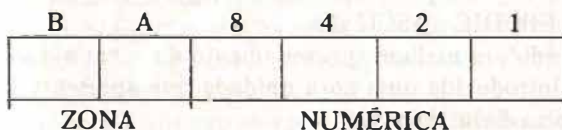
Os quatro bits suficientes para representar os algarismos constituem a parte numérica do dígito. A esses são acrescentados mais dois, usados para representar letras e sinais especiais, que constituem a chamada parte de zona do dígito.

Desta forma, totalizamos 6 bits, o que nos permite representar $2^6 = 64$ símbolos diferentes.

Cada bit recebe uma designação própria para facilitar sua referência.

Os 4 bits da parte numérica são designados, da direita para a esquerda, pelos algarismos 1, 2, 4, 8, que são os valores de posição desses bits. Os dois bits da parte de zona são designados pelas letras A e B, também da direita para a esquerda.

Esquemáticamente temos:



Assim sendo, dentro da memória do computador a representação do número 7.940 passa a ser:

7	9	4	0
000111	001001	000100	000000

Por se tratar da representação de um número, a parte de zona do dígito permanece inalterada.

A título de exemplo apresentamos a seguir a representação dos 10 algarismos e das 26 letras do alfabeto, em código BCD:

BA8421		BA8421	
0	000000	H	111000
1	000001	I	111001
2	000010	J	100001
3	000011	K	100010
4	000100	L	100011
5	000101	M	100100
6	000110	N	100101
7	000111	O	100110
8	001000	P	100111
9	001001	Q	101000
		R	101001
		S	010010
A	110001	T	010011
B	110010	U	010100
C	110011	V	010101
D	110100	W	010110
E	110101	X	010111
F	110110	Y	011000
G	110111	Z	011001

Antes de passarmos a falar dos códigos EBCDIC e ASCII, façamos um parêntese para explicar os códigos de 8 bits, que é o que os códigos EBCDIC e ASCII são.

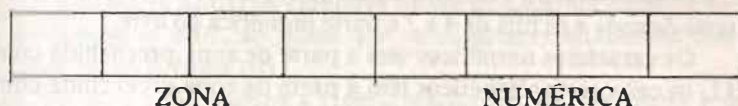
Visando um melhor aproveitamento da memória do computador, foi introduzida uma nova unidade que apresenta vantagens em relação ao dígito de 6 bits.

De fato, um dígito utiliza 6 bits para armazenar um caráter numérico ou alfabético.

Ora, no caso de informações numéricas bastam 4 bits para re-

presentar o algarismo, sendo desnecessários os 2 de zona, que permanecem inutilizados.

Já que 4 bits são suficientes para representar um carácter numérico, por que não aumentar o número de bits de 6 para 8. Assim procedendo teremos a seguinte configuração esquemática:



Desta forma, quando estivermos lidando com informações numéricas, a parte de zona pode ser usada para representar um carácter numérico.

Então, com 8 bits podemos representar 2 algarismos, sendo que da forma anterior seriam necessários 12 bits.

Nasce assim o BYTE, com 8 bits, em substituição ao DÍGITO, com 6 bits.

O byte pode ser usado de duas maneiras que constituem o Formato Descompactado e o Formato Compactado.

No Formato Descompactado o byte armazena um carácter alfanumérico (alfabético ou numérico) e no Formato Compactado ele armazena dois caracteres numéricos.

Em vista da nova forma de armazenamento, cuja unidade básica, o byte, é composta de 8 bits, novos códigos tiveram que ser criados para permitir a representação dos caracteres numéricos e alfanuméricos.

Um desses códigos é uma ampliação do código BCD e recebeu o nome de Extended Binary-Coded Decimal Interchange, sendo conhecido pelas iniciais EBCDIC;

O outro código de 8 bits denomina-se American Standard Code For Information Interchange, mais conhecido por ASCII.

Devemos notar que, graças aos 8 bits que constituem o byte, é possível obter $2^8 = 256$ configurações diferentes. Muito mais do que o número de caracteres usuais.

Graças a este número elevado de possibilidades de representação, os códigos de 8 bits permitem representar números, letras maiúsculas e minúsculas, símbolos especiais, símbolos matemáticos do tipo: $\sqrt{\quad}$, %, \int , \sum , +, -, letras do alfabeto grego como

α , β , π , Δ , Ω , etc. Mesmo assim muitas configurações permanecem sem correspondência.

Código EBCDIC

No código EBCDIC os bits dentro do byte são numerados de 0 a 7, da esquerda para a direita, sendo que os bits de 0 a 3 constituem a parte de zona e os bits de 4 a 7 a parte numérica do byte.

Os caracteres numéricos têm a parte de zona preenchida com 1111, os caracteres alfabéticos têm a parte de zona preenchida com 1100 ou 1101 ou ainda 1110, segundo o caráter a ser representado. Os demais caracteres têm sua parte de zona preenchida com diferentes combinações de 0 e de 1.

A seguir apresentamos os dez algarismos, as 26 letras do alfabeto e alguns símbolos especiais, com as respectivas representações em EBCDIC.

01234567		01234567	
0 11110000	L 11010011
1 11110001	M 11010100
2 11110010	N 11010101
3 11110011	O 11010110
4 11110100	P 11010111
5 11110101	Q 11011000
6 11110110	R 11011001
7 11110111	S 11100010
8 11111000	T 11100011
9 11111001	U 11100100
		V 11100101
A 11000001	W 11100110
B 11000010	X 11100111
C 11000011	Y 11101000
D 11000100	Z 11101001
E 11000101		
F 11000110	+ 01010000
G 11000111	? 11000000
H 11001000	* 01011100
I 11001001	\$ 01011011
J 11010001	/ 01100001
K 11010010	espaço 01000000

Código ASCII

No código ASCII os bits dentro do byte são numerados da seguinte maneira: 7 6 x 5 4 3 2 1.

Portanto os bits 7 6 x 5 formam a parte de zona e os bits 4 3 2 1 formam a parte numérica do byte.

No código ASCII os caracteres numéricos têm na parte de zona a configuração 0101, ao passo que os caracteres alfabéticos têm a configuração 1010 ou 1011.

A seguir mostramos a representação dos 10 algarismos e das 26 letras do alfabeto, no código ASCII:

	76x54321		76x54321
0 01010000	H 10101000
1 01010001	I 10101001
2 01010010	J 10101010
3 01010011	K 10101011
4 01010100	L 10101100
5 01010101	M 10101101
6 01010110	N 10101110
7 01010111	O 10101111
8 01011000	P 10110000
9 01011001	Q 10110001
		R 10110010
		S 10110011
A 10100001	T 10110100
B 10100010	U 10110101
C 10100011	V 10110110
D 10100100	W 10110111
E 10100101	X 10111000
F 10100110	Y 10111001
G 10100111	Z 10111010

Para encerrar, é importante fazer referência ao código hexadecimal que, a rigor, não é propriamente um código interno de armazenamento de dados, mas sim um modo prático de representar dados binários.

De fato, se considerarmos o número binário:

1011001110100111 (A).

Além da dificuldade de leitura, temos o inconveniente de facilitar erros no momento da digitação.

Dividindo este número binário em grupos de 4, da esquerda para a direita, preenchendo com zeros o grupo mais à esquerda, caso não complete 4 códigos. Temos:

1011 0011 1010 0111 (B)

Basta agora substituir cada configuração de 4 bits por seu equivalente hexadecimal e temos:

B3A7 (C)

Como vemos, a forma (C) é uma representação bem condensada da forma (A).

Segue uma tabela de correspondência entre os símbolos hexadecimais e seus correspondentes decimais e em bits:

HEXADECIMAL	DECIMAL	BITS
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
A	10	1010
B	11	1011
C	12	1100
D	13	1101
E	14	1110
F	15	1111

2.5. ORGANIZAÇÃO DE ARQUIVOS

Já vimos que o bit é a menor unidade passível de ser representada num computador.

Vimos também que um conjunto de 8 bits forma um byte, com o qual podemos representar um caráter qualquer (número, letra, etc.).

Para melhor identificação, as informações são tratadas pelo computador em partes chamadas CAMPOS. Estes campos normalmente se referem a conjuntos bem definidos de informações. Por exemplo, no caso de uma referência bibliográfica, nós teríamos o nome do autor como sendo um campo. O título seria outro campo e assim sucessivamente.

A referência completa por sua vez constitui uma unidade completa a que chamamos de REGISTRO.

O conjunto de referências que constitui a coleção ou acervo constitui o que chamamos ARQUIVO.

Assim sendo os dados ou informações têm, no computador, a seguinte hierarquia:

BIT — BYTE — CAMPO — REGISTRO — ARQUIVO.

Modernamente, surgiu o conceito de BANCO DE DADOS, que pode ser enunciado da seguinte forma: "Conjunto de arquivos relacionados e passíveis de serem acessados por vários sistemas de computação".

Portanto, a hierarquia, acrescida de mais um elemento, passa a ser:

BIT — BYTE — CAMPO — REGISTRO — ARQUIVO
— BANCO DE DADOS.

Admitamos a seguinte referência:

TAROUCO, Liane

A teledocumentação no Brasil: situação atual e perspectivas.
Brasília: CENAGRI, 1983

11000001 — 8 bits = 1 byte

Nesta referência temos:

- 1 registro (que é a própria referência);
- 4 campos (autor, título, subtítulo, impressão);

96 bytes (incluindo espaços e sinais de pontuação);
768 bits.

Alguém poderá perguntar: por que 4 campos?

A resposta é a seguinte: a divisão em campos é feita de acordo com o uso que se vai fazer das informações. Por exemplo, se desejarmos processar uma informação como autor, é recomendável definir a autoria como um campo. Se, pelo motivo que for, desejarmos acessar o subtítulo independentemente do título, é recomendável que este subtítulo não pertença ao mesmo campo do título, pois assim seria difícil separar os dois. No exemplo acima foi definido um campo único chamado impressão, onde estão reunidos local de edição, editora e ano de publicação. Se desejássemos trabalhar apenas com o ano, teríamos sérias dificuldades de programação, podendo até ser impossível fazê-lo. Para isto, o ano deveria ter sido definido como um campo isolado.

Disto concluímos que não existem regras fixas para a determinação de campos num registro. Tudo depende dos processamentos desejados.

Por organização de um arquivo entendemos a maneira como os registros são agrupados dentro deste arquivo.

Existem 4 modos de organizar registros num arquivo e a escolha de um ou de outro depende das características próprias do arquivo e do uso que se vai fazer deste.

As características a serem consideradas para organizar um arquivo são basicamente:

VOLATILIDADE — é a quantidade de registros que são adicionados e subtraídos de um arquivo durante um processamento.

Um arquivo com pequena quantidade de adições e subtrações é chamado “arquivo estático”. Caso contrário, “arquivo volátil.”

ATIVIDADE — refere-se à porcentagem de registros do arquivo a serem processados.

Um arquivo que tem pequena porcentagem de registros processados é um arquivo “pouco ativo”. Tendo grande porcentagem trata-se de um arquivo “muito ativo”.

TAMANHO — diz respeito à quantidade de registros que o arquivo contém.

Se tiver muitos registros é um "arquivo grande". Se não, é um "arquivo pequeno".

SUPORTE — por suporte entendemos o dispositivo sobre o qual o arquivo é gravado.

É preciso ter cuidado para não querer organizar um arquivo fora das possibilidades de trabalho do dispositivo.

Tendo em mente estas características podemos organizar um arquivo de forma seqüencial, seqüencial indexada, direta ou particionada.

A seguir veremos, rápida e esquematicamente, as características mais importantes de cada organização.

Organização seqüencial

— Muito simples, fácil de programar e possível em qualquer tipo de dispositivo: fita, disco, cartões, etc.

— Os registros estão em seqüência física, sem nenhum relacionamento lógico.

— Para localizar um determinado registro é preciso passar por todos os anteriores.

— A retirada ou o acréscimo de um registro implica a necessidade de recriar o arquivo.

Organização seqüencial indexada

— Os registros estão em seqüência física, mas existem índices que permitem localizar os registros diretamente sem ter que passar por todos os anteriores. Como consequência, este tipo de organização pode ser usada somente em dispositivos de acesso direto.

— Neste tipo de organização, os registros devem ser ordenados previamente em seqüência crescente de chave. (Chave é uma parte do próprio registro usada como elemento identificador do mesmo, diferenciando-o dos demais. Ex: número de tombo.)

— Permite a inserção e a retirada de registros sem necessidade de recriação imediata do arquivo.

— Um arquivo seqüencial indexado é constituído de três áreas:

- Área primária, contendo os registros propriamente ditos.

- Área de índices, contendo apontadores que permitem localizar os registros.
- Área de overflow, contendo as inclusões de registros que não couberam na área primária.

Organização direta

- Os registros são acessados diretamente sem necessidade de passar pelos anteriores e sem necessidade de índices.
- Permite a inserção e a eliminação de registros sem necessidade de recriação do arquivo.
- Possível somente em dispositivos de acesso direto.
- Usada para arquivos com características que não permitem outro tipo de organização ou para arquivos em que a localização de um registro deve ser a mais rápida possível.
- Nesta organização há dois tipos de endereçamento:
 - Endereçamento direto, onde a chave do registro é o próprio endereço.
 - Endereçamento indireto, onde a relação entre a chave e o endereço não é imediata, dependendo de um cálculo preestabelecido.

Organização particionada

- Reúne características da organização seqüencial e da seqüencial indexada.
- Trata-se de um conjunto de arquivos independentes, organizados seqüencialmente, chamados membros. Os registros dentro dos membros são organizados seqüencialmente.
- Cada membro está relacionado, junto com seu endereço, numa parte do arquivo denominada “diretório”.
- A atualização não pode ser feita no próprio membro a ser modificado, mas sim numa outra área do arquivo (desde que haja espaço), permanecendo a área inicial inaproveitável até a próxima reorganização.
- Organização apropriada para armazenar dados seqüenciais tais como: tabelas, programas, etc.

2.6. FLUXOGRAMAS

Freqüentemente, durante a fase de planejamento de uma atividade qualquer, é preciso fazer uma descrição da seqüência de operações necessárias à realização daquela atividade.

Para uma melhor visualização, essa descrição pode ser feita sob forma de um gráfico onde cada operação elementar é representada por um símbolo convencional.

Em processamento de dados é muito comum usar-se um tipo de descrição gráfica denominada FLUXOGRAMA.

Existem dois tipos de fluxogramas: Fluxograma de Sistema e Fluxograma de Programação.

Tanto um como o outro utilizam símbolos padronizados que permitem sua interpretação de maneira idêntica por qualquer pessoa iniciada.

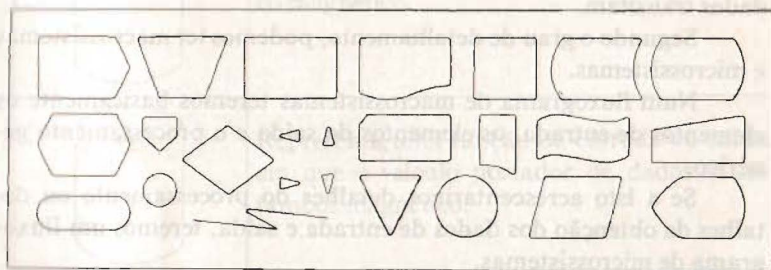
Esta padronização foi estabelecida em 1966 pelo USASI — United States of America Standard Institute.

Os símbolos padronizados pelo USASI podem ser complementados por outros propostos pela ISO — International Standard Organization.

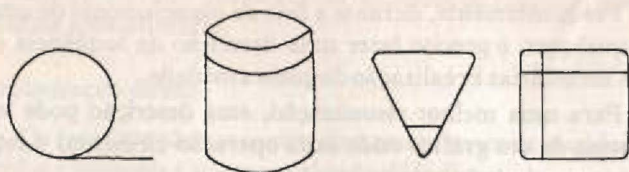
O tamanho de um símbolo pode variar desde que permaneça constante a relação entre a altura e a largura.

É facultativo escrever dentro de um símbolo.

Existem “régua” apropriadas contendo todos os símbolos usados nos fluxogramas. A seguir é apresentado um exemplo deste tipo de régua.



Alguns símbolos são obtidos a partir da junção ou da combinação de mais de um símbolo da régua. Ex:

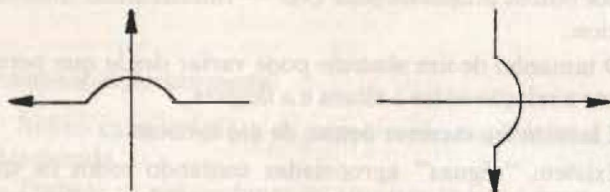


Os símbolos são ligados entre si por linhas que indicam o fluxo de execução.

A direção normal de leitura do fluxograma é de cima para baixo e da esquerda para a direita. Nos casos em que o sentido for diferente ou for desejada maior clareza, é aconselhado o uso de setas.

Toda vez que um fluxo tiver que ser interrompido, serão usados conectores nas interrupções.

No cruzamento de duas linhas é aconselhável evitar o cruzamento direto, usando linhas que se circundem, evitando confusão. Ex:



Fluxogramas de sistema

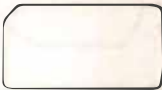




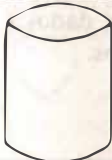
Descrevem o processamento a ser efetuado, mostrando as etapas, os dados iniciais e finais, bem como os veículos em que estes dados transitam.

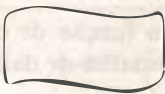



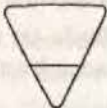

Segundo o grau de detalhamento, podemos ter macrossistemas e microssistemas.







Num fluxograma de macrossistemas teremos basicamente os elementos de entrada, os elementos de saída e o processamento genérico.

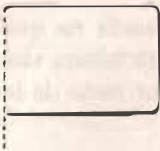

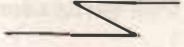
Se a isto acrescentarmos detalhes do processamento ou detalhes da obtenção dos dados de entrada e saída, teremos um fluxograma de microssistemas.

Apresentamos a seguir os símbolos usados em fluxogramas de sistemas e os respectivos significados:

	<p>Este símbolo representa uma função de entrada e saída, cujo veículo portador de dados é o cartão.</p>
	<p>Representa início, fim ou interrupção.</p>
	<p>Qualquer operação manual sem auxílio de nenhum tipo de máquina.</p>
	<p>Este símbolo representa uma função de entrada ou saída, cujo veículo portador de dados é a fita magnética.</p>
	<p>Representa uma função de entrada ou saída em que o veículo portador de dados é um disco magnético.</p>
	<p>Representa uma função de entrada ou saída em que o veículo portador de dados é um tambor magnético.</p>

	<p>Representa uma função de entrada ou saída onde o veículo portador de dados é uma fita de papel perfurada.</p>
	<p>Representa uma função de entrada ou saída onde o veículo portador de dados é um documento.</p>
	<p>Representa uma entrada de dados manual, durante o processamento, por meio de teclados "on-line".</p>
	<p>Representa uma função de entrada ou saída que utiliza memórias auxiliares de acesso direto às quais o sistema tenha acesso: disco, tambor ou célula.</p>
	<p>Representa a armazenagem externa de informações, qualquer que seja este meio: arquivo, gaveta, etc.</p>
	<p>Representa uma função de entrada e saída, na qual o veículo portador de dados é uma memória de núcleos magnéticos.</p>

	<p>Representa uma função de saída na qual a informação é apresentada para leitura visual, durante o processamento, por meio de indicadores "on-line".</p>
	<p>Representa uma função de processamento eletrônico, ou seja, a execução de uma ou mais operações.</p>
	<p>Representa uma operação em que se utiliza um teclado "off-line": ex: uma perfuradora de cartões.</p>
	<p>Representa uma fita de máquina de somar.</p>
	<p>Representa uma entrada ou uma saída de uma para outra parte do fluxograma.</p>
	<p>Representa uma entrada ou uma saída de uma página a outra do fluxograma.</p>

	<p>Representa a inclusão de uma explicação adicional ao fluxograma.</p>
	<p>Indicam a direção de fluxo.</p>
	<p>Indica a transmissão de dados via teleprocessamento.</p>


Vide, na página seguinte, exemplo de fluxograma de sistemas.

Fluxogramas de programação

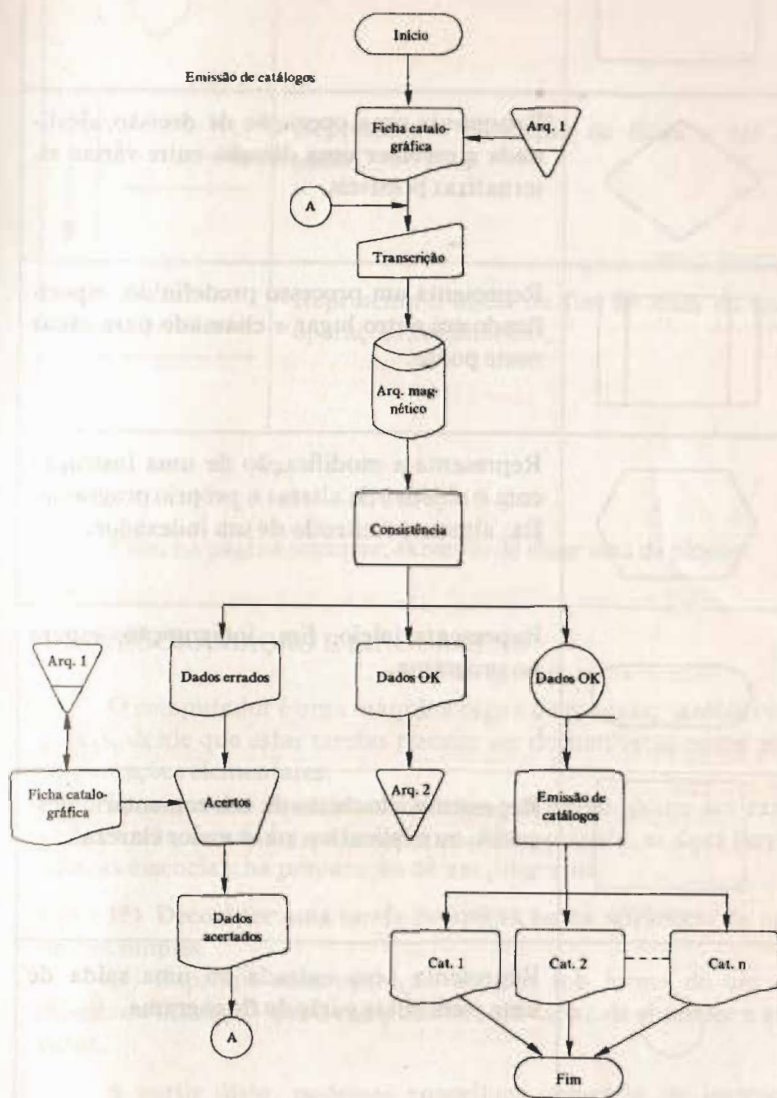
O fluxograma de programação, mais conhecido como diagrama de blocos, tem o objetivo de descrever graficamente a solução de um problema.





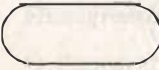


Geralmente, cada etapa da solução lógica do problema é representada por um símbolo cujo formato identifica o tipo de processo envolvido.


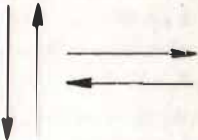

Estes símbolos, apresentados a seguir, recebem o nome genérico de blocos. Daí o nome de diagrama de blocos.

	<p>Este símbolo representa uma função de entrada ou saída.</p>
---	--

Fluxograma de sistemas



	<p>Representa uma função de processamento.</p>
	<p>Representa uma operação de decisão, destinada a escolher uma direção entre várias alternativas possíveis.</p>
	<p>Representa um processo predefinido, especificado em outro lugar e chamado para atuar neste ponto.</p>
	<p>Representa a modificação de uma instrução com o objetivo de alterar o próprio programa. Ex: alterar o conteúdo de um indexador.</p>
	<p>Representa início, fim, interrupção, espera no programa.</p>
	<p>Representa a inclusão de um comentário descritivo ou explicativo, para maior clareza.</p>
	<p>Representa uma entrada ou uma saída de uma para outra parte do fluxograma.</p>

	Representa um conector de páginas.
	Representa a indicação de fluxo a ser seguido.
	Representa o início ou fim de duas ou mais operações simultâneas.

Vide, na página seguinte, exemplo de diagrama de blocos.

2.7. PROGRAMAÇÃO E LINGUAGENS

O computador é uma máquina capaz de executar tarefas complexas, desde que estas tarefas possam ser decompostas numa série de operações elementares.

Estas tarefas elementares por sua vez devem poder ser explicadas à máquina de forma inequívoca. Eis, portanto, as duas tarefas básicas essenciais na preparação de um programa:

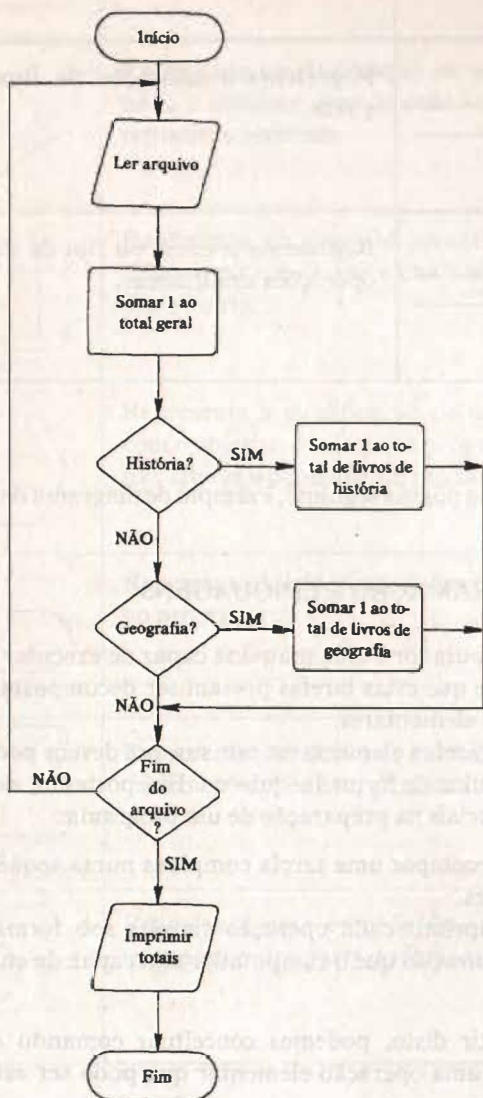
1^a) Decompor uma tarefa complexa numa sequência de operações simples.

2^a) Expressar cada operação simples sob forma de um *comando* ou *instrução* que o computador seja capaz de entender e executar.

A partir disto, podemos conceituar comando ou instrução como sendo uma operação elementar que pode ser executada pelo computador.

Diagrama de blocos

Numa coleção desejamos conhecer o total de livros de história, o total de livros de geografia e o total geral da coleção.



Cada modelo de computador dispõe de um determinado conjunto de instruções que constitui seu repertório de instruções.

Este repertório pode variar de algumas dezenas até mais de uma centena.

De modo geral, as instruções são compostas por VERBOS, que determinam alguma função a ser executada, OPERANDOS, que são os campos a serem trabalhados, e CLÁUSULAS, que complementam a função a ser executada.

Ao conjunto de instruções necessárias para realizar uma determinada tarefa dá-se o nome de PROGRAMA.

O programa é o roteiro que orienta o computador, passo a passo, mostrando-lhe a seqüência das operações necessárias à execução de uma tarefa de processamento eletrônico de dados.

Quanto mais complexa for a tarefa, maior será o número de instruções do programa, que pode alcançar alguns milhares.

É de se notar, entretanto, que nem todas as instruções são executadas em todos os casos, assim como existem instruções que podem ser executadas inúmeras vezes.

Como sabemos, o computador trabalha em binário e a comunicação entre ele e o homem teria que ser efetuada utilizando esta linguagem.

Isto é extremamente difícil e trabalhoso, por isso foram desenvolvidas linguagens que facilitem esta comunicação.

Criaram-se, portanto, linguagens mais facilmente manipuláveis pelos programadores, o que implicou a necessidade de "tradutores", capazes de transformar estas linguagens na linguagem do computador, conhecida como linguagem de máquina.

Estes tradutores recebem o nome de COMPILADORES.

Desta forma temos a seguinte seqüência:

Um programador escreve seu programa numa linguagem de fácil utilização, gerando o que se chama PROGRAMA FONTE. Este é dado ao computador, o qual, por meio do compilador, transforma-o em linguagem de máquina, originando o chamado PROGRAMA OBJETO, que é executável.

Existem várias linguagens de programação e a escolha de uma ou de outra depende basicamente das características de cada uma delas.

As linguagens de programação podem ser agrupadas em três grandes categorias:

Linguagens de máquina,
Linguagens de baixo nível e
Linguagens de alto nível.

Linguagens de máquina

São linguagens próprias de cada máquina em particular, constituídas por seqüências de bits.

Exemplo de instrução em linguagem de máquina:

00011000 00011101

São as únicas linguagens que o computador “entende”, dispensando, portanto, o uso de compiladores.

Todas as outras linguagens devem ser convertidas nesta.

Linguagens de baixo nível

São linguagens voltadas para a máquina. Isto quer dizer que, ao programar usando uma linguagem de baixo nível, é preciso conhecer as características da máquina onde o programa vai ser executado.

Difícilmente um programa escrito numa linguagem de baixo nível, para um certo modelo de computador, pode ser executado em modelo diferente.

O número de instruções, numa linguagem simbólica de baixo nível, é aproximadamente igual ao número de instruções produzidas, pelo compilador, em linguagem de máquina.

Exemplo de instrução em linguagem de baixo nível:

ST 3,AREA.

Entre as vantagens das linguagens de baixo nível podemos citar:

- compiladores menos complexos, usando menos memória principal;

- por estar mais próxima da máquina, a linguagem de baixo nível permite o uso de artifícios que seriam impossíveis com as linguagens de alto nível;

- menor espaço de memória ocupado pelo próprio programa, pelo fato de ser específico e não realizar operações de maneira

generalizada, como fazem os programas escritos em linguagens de alto nível.

O ASSEMBLER é um exemplo de linguagem de baixo nível.

Linguagens de alto nível

São linguagens voltadas para o problema. Isto quer dizer que não é necessário conhecer as características do computador que irá executar o programa.

De modo geral, um programa escrito numa linguagem de alto nível pode ser executado em diferentes modelos de computadores, bastando pequenas modificações.

O número de instruções numa linguagem de alto nível é muito menor do que o número de instruções geradas após a compilação. Isto porque o compilador tem a capacidade de "encaixar" no programa certas rotinas já existentes e padronizadas.

Exemplo de instrução em linguagem de alto nível:

ADD 3 TO 4 GIVING RESULTADO.

A maior vantagem das linguagens de alto nível é a extrema facilidade de uso.

São exemplos de linguagens de alto nível: COBOL, FORTRAN, PL/I, PASCAL, ADA.

2.7.1. Software aplicativo e software básico

Como foi dito no início, toda parte de programação de um computador recebe o nome genérico de software (em contraposição a hardware que designa a própria máquina).

O software, que instrui a máquina sobre o que deve ser feito e como deve ser feito, pode ser dividido em dois grandes grupos: software aplicativo e software básico.

Software aplicativo

São programas específicos para resolver problemas específicos. São escritos, geralmente, em linguagens de alto nível, pelos próprios usuários do computador, ou por software-houses (firmas que desenvolvem programas para finalidades específicas como, por

exemplo: controle de estoque, contas a pagar, folha de pagamento e contabilidade).

Software básico

São programas fornecidos pelo próprio fabricante da máquina (via de regra), geralmente escritos em linguagem de máquina (que os dispensa de serem traduzidos) e que, segundo sua finalidade, podem ser divididos em:

- Sistemas operacionais: cuja finalidade é controlar o funcionamento dos demais programas. Ex: OS-VS, DOS-VM, CP/M, etc.

- Compiladores: que têm por finalidade traduzir e interpretar os programas escritos em outras linguagens que não a linguagem de máquina.

- Editores: para entender melhor a função deste programa, é preciso saber que um determinado programa aplicativo pode ser desenvolvido em módulos ou partes, os quais podem inclusive ser codificados em linguagens diferentes. Cabe aos editores juntar estes módulos (após terem sido traduzidos pelos compiladores), resolvendo todas as referências externas de cada um e as ligações entre eles, transformando-os num único módulo executável.

- Utilitários: embora possam ser escritos também pelos usuários (e muitos o são), geralmente um conjunto básico deles é fornecido pelo fabricante da máquina. São programas utilizados para efetuar tarefas comuns no dia-a-dia, como por exemplo copiar arquivos, classificar arquivos, etc. São programas bastante genéricos aos quais o usuário fornece certos parâmetros que os tornam específicos para determinado uso. Terminada a tarefa, eles podem receber outros parâmetros tornando-se aptos a efetuar outra tarefa e assim sucessivamente.

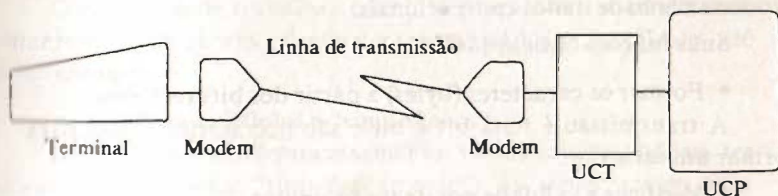
2.8. TELEPROCESSAMENTO

Teleprocessamento ou processamento remoto é a soma dos conceitos de processamento eletrônico de dados e de telefonia.

De fato, utilizando os recursos da telefonia é possível processar dados a partir do lugar de sua geração, sem necessidade de transportá-los até o local onde se encontra o complexo central ou CPU.

Isto permite a um caixa de banco verificar imediatamente o saldo de um cliente, a um pesquisador efetuar seus cálculos complexos sem sair do seu laboratório ou acessar bancos de dados internacionais, e assim sucessivamente.

Esquemáticamente, um sistema de teleprocessamento pode ser apresentado da seguinte forma:

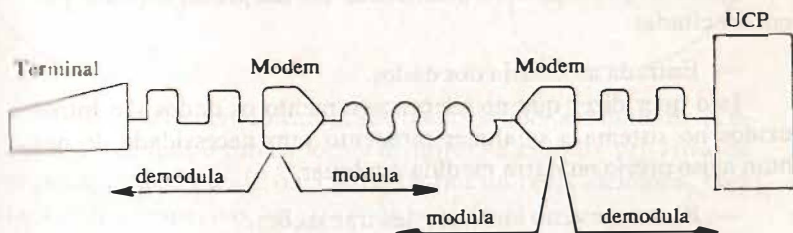


onde:

— Terminal é um conjunto de um ou mais dispositivos de entrada e/ou saída, orientados para comunicação. Ex: caixa eletrônico de banco 24 horas.

— Modem é o elemento de ligação entre o terminal ou a UCP e a linha de transmissão. Tem por função modular o sinal que parte (do terminal ou da UCP) e demodular o sinal que chega (do terminal ou da UCP).

Cabe aos modems, portanto, transformar os sinais digitais originados pelo computador (terminal ou UCP) em sinais analógicos transmissíveis e vice versa.



— Linha de transmissão é o nome genérico atribuído a cabos ou canais em geral, sobre os quais veiculam as informações num sistema de TP.

Podem ser fios telefônicos, cabos submarinos, ou mesmo microondas.

— Unidade de Controle de Transmissão (UCT), usada para compatibilizar os sistemas de teleprocessamento com os sistemas de processamento de dados convencionais.

Suas funções básicas são:

- Formar os caracteres (bytes) a partir dos bits recebidos.

A transmissão é feita bit a bit e são necessários 8 bits para formar um caractere.

- Verificar a validade dos dados recebidos.

Quando uma mensagem é transmitida, um caráter de verificação é gerado pelo dispositivo transmissor.

Quando a mensagem é recebida, é gerado um caráter de verificação pelo dispositivo receptor.

O caráter transmitido é comparado com o caráter recebido e é gerada uma informação indicando se a mensagem é correta ou não.

- Inserir e retirar estes caracteres de controle.

Os caracteres de controle são inúteis para o processamento e são incluídos no texto que sai, sendo retirados do texto que chega.

2.8.1. Características do Teleprocessamento

Entre as principais características do teleprocessamento podem ser citadas:

- Entrada arbitrária dos dados.

Isto quer dizer que no teleprocessamento os dados são introduzidos no sistema a qualquer momento sem necessidade de nenhum aviso prévio ou outra medida qualquer.

- Processamento imediato das transações.

Uma vez dada a entrada dos dados no sistema, este efetua o processamento imediatamente, fornecendo as respostas.

- Tempo de resposta rápido.

A velocidade no tempo de resposta pode variar um pouco dependendo de fatores como: tipo de equipamento utilizado, velocidade das linhas de transmissão usadas, volume de trabalho a ser realizado pela UCP, quantidade de usuários a serem atendidos pelo sistema, etc.

- Comprimento das mensagens variável.

O usuário pode trabalhar com informações dos mais diversos tamanhos, desde alguns números a serem usados para cálculo, até textos enormes.

- Usuários múltiplos e “simultâneos”.

Num sistema de teleprocessamento, vários usuários podem ter acesso ao computador “simultaneamente”. As aspas servem para chamar a atenção para o fato de esta simultaneidade ser aparente.

De fato, o computador realiza suas funções atendendo aos usuários concorrentemente (um após o outro). Entretanto, a velocidade de processamento é tão grande que dá a impressão de que todos os usuários estão sendo atendidos ao mesmo tempo.

2.8.2. Conceitos básicos em Teleprocessamento

- Circuitos: são conexões entre dois pontos a partir da composição de equipamentos e meios que permitam o transporte da informação codificada.



- Meios de comunicação: são recursos que possibilitam a comunicação de dois ou mais locais distintos. Ex: linha privativa de comunicação de dados (2 ou 3 pares de fios da rede telefônica, com características especiais para comunicação de dados).

- Tipos de transmissão: são as várias modalidades com que é possível transmitir mensagens:

Simplex: transmissão num único sentido. Exemplo: televisão.
Semiduplex: transmissão nos dois sentidos, mas não simultaneamente. Exemplo: rádio-amador.

Duplex: transmissão nos dois sentidos, simultaneamente. Exemplo: telefone.

Síncrona: transmissão de informações em blocos, com caracteres indicando início e fim de bloco.

Assíncrona: transmissão sem limitações de tamanho da mensagem, que pode ser iniciada a qualquer tempo.

— Velocidade de transmissão: é a quantidade de bits que pode ser transmitida por unidade de tempo (geralmente segundo).

As mais usuais são: 1.200 BPS (Bits Por Segundo), 2.400 BPS, 4.800 BPS, 9.600 BPS.

Entretanto, existem também velocidades como 300 BPS, 19.200 BPS, 48.000 BPS..., menos usuais.

2.9. MICROCOMPUTADORES

O avanço da tecnologia tem propiciado o aparecimento de circuitos integrados cada vez menores e mais rápidos.

Sem dúvida, o CHIP causou grande impacto no mundo da informática, pois, graças à compactação de milhares de componentes como resistências, diodos e transistores, foi possível a construção de poderosos e completos sistemas de computação numa única pastilha de tamanho não maior que um centímetro quadrado.

Estes chips ou pastilhas viabilizaram a construção de microcomputadores, dos quais são os componentes mais importantes.

Em geral um microcomputador se compõe de:

— TELA ou display de vídeo onde são exibidos os resultados e os valores digitados pelo operador.

— TECLADO com teclas numéricas, alfabéticas, de sinais especiais e de funções.

— PLACAS de circuitos eletrônicos que contêm a UCP e a memória.

— DRIVERS ou unidades de acionamento de discos magnéticos.

— IMPRESSORA que permite obter respostas impressas em papel.

As memórias de circuitos integrados podem ser:

ROM (Read Only Memory) — Já vem gravada pelo fabricante com dados e instruções. Por isso, nada pode ser gravado nela. Apenas lido.

RAM (Random Access Memory) — O usuário pode tanto ler como gravar dados nesta memória. É a memória de trabalho.

A memória do tipo ROM pode, por sua vez, ser dividida em dois grupos: PROM e EPROM.

PROM (Programable Read Only Memory) — cujo conteúdo é gravado pelo fabricante não podendo mais ser alterado.

EPROM (Erasable Programable Read Only Memory) — cujo conteúdo pode ser apagado pelo fabricante (através de um banho de raios ultravioleta) e reprogramado pelo fabricante, de acordo com as necessidades do usuário.

Os discos magnéticos de um microcomputador têm superfícies magnetizadas onde é possível armazenar dados.

Podem ser de dois tipos:

Discos flexíveis e removíveis (floppy disks), conhecidos como disquetes.

São discos de plástico magnetizado, acondicionados dentro de um envelope de papelão. Podem ser retirados da unidade e guardados para uso posterior.

Seu custo é baixo, porém a imprecisão, devido ao desgaste por manuseio, é grande.

Discos rígidos fixos tipo winchester.

Possuem maior capacidade de armazenamento e maior velocidade de operação, mas não podem ser retirados da unidade.

Quase não sofrem desgaste, mas seu custo é alto.

As impressoras podem ser classificadas basicamente em dois grupos:

MATRICIAIS e DE IMPACTO.

As matriciais são caracterizadas por um tipo de impressão que tem como base uma matriz de pontos.

Pertencem a este grupo os modelos Emilia, Mônica Plus e Alice.

As impressoras de impacto utilizam caracteres já formatados e colocados em dispositivos especiais (como esferas, margaridas, etc.) que são impressos pelo impacto causado por um dispositivo geralmente chamado de martelo.

Uma das impressoras de impacto mais conhecida é a Daysy Wheel, onde os caracteres são colocados no dispositivo margarida.

2.10. DESENVOLVIMENTO DE SISTEMAS

Automatizar os processos de uma biblioteca ou de um centro de documentação é uma tarefa que apresenta um grau de dificuldade considerável.

Primeiramente porque este é um tipo de automação que foge aos padrões costumeiros de automação como: controle de estoques, folhas de pagamento, contas a pagar e a receber, contabilidade, faturamento, contas correntes, e assim por diante.

Isto quer dizer que não são muitos os profissionais de processamento eletrônico de dados que têm desembaraço neste tipo de aplicação.

Em segundo lugar porque as biblioteconômicas e as documentárias são tarefas diferenciadas e detalhadas a ponto de dificultarem soluções globais e lineares. Isto sem contar que até pouco tempo atrás (e em certos casos até hoje) os bibliotecários tinham certa dificuldade em dialogar com um analista de sistemas de computador, tanto para expor-lhe suas necessidades como para ouvir deste eventuais colocações.

Hoje, de modo geral, a situação melhorou quanto a este último aspecto. Em parte porque o computador já invadiu o cotidiano de cada um de nós e em parte porque a própria formação acadêmica vem treinando os alunos de biblioteconomia a, pelo menos, não ter medo dos computadores nem do "computês" falado, nem sempre com propriedade, pelo *homo informaticus*.

Entretanto, apesar das diferenças de posturas e dos resultados obtidos, a automação de serviços biblioteconômicos e/ou documentários também implica a adoção de uma metodologia apropriada para o desenvolvimento e a implantação destes sistemas.

Isto quer dizer que existem etapas a serem percorridas entre o desejo ou a necessidade de se automatizar um sistema e sua implantação.

O conhecimento dessas etapas, ainda que de forma superficial, pode contribuir para um melhor entendimento do mecanismo, além de permitir um diálogo mais efetivo entre profissionais de áreas diferentes que se juntam para alcançar um objetivo comum.

Isto quando o próprio bibliotecário não se torna, de certa forma, um analista de sistemas para resolver seus problemas de automação.

Neste caso, este conhecimento se torna imperioso e aí estas linhas podem não ser suficientes, ainda que úteis.

Existem muitas metodologias para desenvolver sistemas. Algumas sofisticadas, outras mais simples, mas todas visando racionalizar o trabalho, evitando gastos desnecessários de recursos, perda de tempo e aquele "vai e vem" típico de quem "já ouviu falar no assunto" e "tem uma leve idéia" dos resultados que deseja obter.

Assim, vamos apresentar a seguir, não propriamente uma metodologia, mas um roteiro que possa auxiliar os não iniciados em suas árduas andanças pelos caminhos do desenvolvimento de sistemas.

Como ponto de partida é preciso saber exatamente o que se quer. Isto, partindo da premissa de que a automação já foi considerada necessária e irreversível.

Para se saber o que se deseja como resultado da automação, é interessante estudar cuidadosamente o sistema vigente, para descobrir-lhe os aspectos positivos e negativos (seria absurdo automatizar tarefas que não correspondem às expectativas).

Nesta fase de estudo detalhado da situação atual, normalmente são detectados problemas com os quais se conviveu durante muito tempo sem perceber-lhes a existência ou, pelo menos, a real dimensão.

Analisar detalhadamente a situação atual implica:

- observar o que está acontecendo, como está acontecendo, o grau de satisfação dos envolvidos no processo, etc.;
- entrevistar funcionários e usuários, tentando tirar dos mesmos todas as informações que possam vir a ser úteis, bem como considerar críticas e sugestões;

— ler documentos pertinentes à tarefa a ser automatizada, buscando subsídios, quer na documentação interna, quer na literatura existente a respeito;

— estudar processos iguais ou semelhantes que já tenham passado pelo processo de automação, verificando seu funcionamento, seus resultados e trocando idéias com as pessoas que participaram da automação, para ter uma dimensão aproximada da problemática, tentando evitar erros já cometidos por outros e buscando repetir ou mesmo melhorar o que de acertado foi feito.

O resultado desta primeira etapa vai fornecer os elementos para projetar o novo sistema automatizado.

É interessante que primeiramente se faça um projeto geral definindo os resultados desejados e as informações a serem geradas.

Deve ser elaborado um fluxograma geral que mostre os procedimentos principais e o fluxo das informações.

Dependendo das circunstâncias, este fluxograma geral deve ser discutido com os superiores hierárquicos, aos quais cabe, normalmente, a aprovação e a liberação dos recursos financeiros.

Feito este projeto geral e aprovado, o próximo passo é detalhar todos os procedimentos a nível operacional.

Nesta etapa, cada bloco do fluxograma geral deve ser pormenorizado, gerando fluxogramas bem detalhados onde constem as origens dos dados, os veículos sobre os quais estes dados são registrados, os diversos processamentos aos quais eles serão submetidos e as saídas previstas.

Ao elaborar estes fluxogramas detalhados deve se ter em mente também a integração de tarefas semelhantes, bem como a eliminação de tarefas desnecessárias. Além disso é muito importante estabelecer controles eficientes e prever caminhos alternativos nos pontos-chave do sistema.

É nesta fase que são elaborados os *lay-outs* das planilhas de entrada de dados (eletrônicas ou em papel), onde é definido cada campo, sua localização, seu tamanho, seu conteúdo, etc.

Todas as saídas devem ser definidas cuidadosamente, tendo em vista sua utilização posterior; as saídas impressas, como por exemplo os catálogos, devem mostrar claramente a disposição e o tipo de informações ao longo da listagem, indicando eventuais títulos, subtítulos, numeração de páginas, notas de rodapé, impres-

são em colunas ou não, etc.; as saídas em meio magnético (usadas para processamentos posteriores), também devem ter seu *lay-out* bem definido, de modo a permitir que as informações registradas possam vir a ser processadas por outros programas, para outros fins.

Além dos *lay-outs* de entrada e saída, devem ser previstos todos os processamentos necessários (consistências, ordenações, emissões de listas de trabalho, relatórios estatísticos, gravações, etc.), desde o início até o fim do processo.

Cada programa deve ser descrito com detalhes suficientes que permitam, a quem for programar, saber exatamente o que ele deve fazer.

A fase seguinte é justamente aquela onde os programas são elaborados a partir das especificações fornecidas.

Estes programas, independentemente de quem os desenvolva e da linguagem utilizada, devem ter sua lógica definida em diagramas de blocos e, após prontos, devem ser cuidadosamente testados, para verificar se fazem aquilo para que foram desenvolvidos.

Uma vez prontos e testados os programas, é possível planejar a fase de implantação do novo sistema, tomando todas as medidas necessárias para que esta implantação aconteça sem imprevistos e sem traumas.

É neste momento que deve ser preparado um programa de treinamento de todo o pessoal que irá operar o sistema e, se for necessário, dos próprios usuários aos quais os produtos do sistema se destinam.

Falando em pessoal, é bom lembrar que deve haver uma postura no sentido de motivar os envolvidos no processo, pois reações negativas, por parte destes, podem criar obstáculos ou mesmo emperrar o andamento do trabalho. Afinal, por mais automatizado que o sistema seja, haverá sempre pessoas a alimentá-lo, a regulá-lo e a usufruir de seus resultados.

Assim que estiver tudo pronto e todo o pessoal treinado, será efetuada a implantação-teste, ou seja, o sistema como um todo é colocado em operação e seu funcionamento é acompanhado de perto, visando assegurar-se de que está funcionando segundo o planejado e detectando e corrigindo possíveis erros ou omissões.

Acompanhar o sistema inclui, evidentemente, verificar a participação do pessoal treinado.

3

INFORMÁTICA APLICADA

"Mais vale uma ficha catalográfica na mão do que duas bases de dados voando."

3.1. AUTOMAÇÃO DOS SERVIÇOS DE UMA BIBLIOTECA

Ao dizer "automação dos serviços de uma biblioteca", queremos dizer automação dos processos técnicos dessa biblioteca. Basicamente, aquisição, emissão de catálogos e/ou índices e circulação.

Freqüentemente esta automação é confundida com criação e exploração de bases de dados contendo o acervo da biblioteca.

Trata-se de coisas distintas, com enfoques e resultados distintos, envolvendo softwares diferentes.

É bem verdade que há softwares que contemplam ambas aplicações, mas o importante é chamar a atenção sobre esta diferença, pois uma confusão neste sentido pode acarretar problemas sérios quando da escolha de um equipamento ou da adoção de um programa.

Assim sendo, neste capítulo faremos algumas considerações sobre automação dos processos de aquisição, emissão de catálogos e controle de empréstimos, deixando para o capítulo seguinte os aspectos voltados à chamada documentação automatizada.

A divisão é meramente didática, não havendo o mínimo interesse em reacender a já batida e cansativa polêmica que visa esta-

belecer os limites de atuação entre uma biblioteca e um centro de documentação ou de informações.

Quantas bibliotecas exercem funções ditas documentárias, sem com isso entrarem em conflito com seu papel e, pelo contrário, exercendo-o melhor.

Em compensação há bibliotecas que nem desconfiam o acervo que têm, nem o papel que lhes cabe na sociedade.

Por isso pode haver diferença entre a biblioteconomia e a documentação, mas não entre biblioteca e centro de documentação.

Uma vez externada nossa opinião, deixemos a discussão sobre as fronteiras aos filósofos da informação e vamos tecer algumas considerações sobre automação de bibliotecas.

Quando uma biblioteca opta por automatizar seus processos técnicos, uma série de estudos devem preceder o início dos trabalhos.

Estes estudos devem ser conduzidos no sentido de se verificar a famosa relação custo \times benefício, considerando que nada pode ser considerado caro ou barato se não for medido, e muito bem medido, o grau de benefício que pode ser obtido como resultado.

Assim sendo, devem ser formuladas perguntas do tipo:

— A biblioteca tem condições de utilizar um computador já existente na empresa, na instituição, na escola, etc., à qual pertence, ou será necessário adquirir um?

— A magnitude e o tipo de acervo justificam o esforço e o custo?

— A quantidade de usuários, reais e potenciais, bem como o tipo de consultas, também os justificam?

— Existem projetos ou perspectivas de a biblioteca passar a fazer parte de uma rede de bibliotecas?

— Pretende-se automatizar alguma tarefa específica ou deve-se pensar numa automação total e integrada?

— A automação pretendida pode ser conseguida com a utilização de microcomputadores ou requer computadores de grande porte?

— Existe no mercado de softwares algum produto que atenda às necessidades ou será preciso desenvolvê-lo?

— Se for preciso desenvolver o software, dispõe-se de pessoal capacitado ou deve ser contratado (independente do vínculo empregatício)?

Estas e muitas outras perguntas devem ser feitas, tendo em mente que todo e qualquer processo de automação, em maior ou menor escala, além de ser demorado e oneroso, ainda altera os procedimentos de trabalho, as formas de controle e regulação, criando uma nova distribuição de poder e de responsabilidades.

Uma vez feita a opção pela automação, seja como resultado de um exame criterioso, seja por razões políticas ou de *status*, que não cabe discutir aqui, é preciso operacionalizar esta automação.

A seguir, serão vistos aspectos genéricos sobre a automação de uma biblioteca, evitando entrar em particularidades referentes ao tipo de biblioteca — pública, escolar, empresarial, de livre acesso, etc. —, tipo de acervo — livros, periódicos, microformas, audiovisuais, etc. — ou ainda tipo de equipamento — microcomputadores, computadores de grande porte, dedicados, compartilhados, IBM ou Burroughs ou outra marca qualquer.

Preocupar-nos-emos mais com as posturas e com as implicações do que propriamente com os aspectos técnicos, existentes na literatura especializada.

3.1.1. Aquisição

O processo de aquisição automatizada implica a necessidade de um arquivo magnético, bem atualizado, contendo os dados dos fornecedores aos quais são enviados os pedidos.

Este processo tem como ponto de partida a solicitação de aquisição do material.

Os dados disponíveis neste momento são gravados num arquivo magnético e permanecerão aguardando a chegada do material ou o cancelamento do pedido, se isto se fizer necessário, seja pelo motivo que for.

O computador poderá ser instruído no sentido de, após um tempo previamente estabelecido, emitir uma reclamação pela não chegada do material.

Isto é feito fornecendo certos parâmetros que permitam ao computador determinar o tempo transcorrido desde o dia da gravação dos dados de solicitação.

Com estes dados mais os dados retirados do arquivo de forne-

cedores, o computador tem condições de imprimir uma reclamação-padrão, completando-a com dados específicos.

A chegada do material vai desencadear uma série de atividades.

a. Os dados gravados no momento do pedido serão completados (corrigindo as eventuais imprecisões) com dados retirados do próprio material.

b. A rotina de reclamação será desativada.

c. O processo de pagamento pode ser agilizado desde que se tenha previsto um arquivo com moedas estrangeiras, de forma que o computador possa efetuar o cálculo da quantia a pagar, deduzir este valor do total disponível para compra, mantendo enfim toda a parte contábil atualizada.

Considerando a diferença de características, é preciso prever os vários tipos de aquisição que podem ser efetuados pela biblioteca.

Apenas como lembrete citamos: pedido avulso, assinatura, pedido permanente, anuidade, grátis, permuta, material adquirido antes da automação e que de alguma forma deve entrar no sistema, etc.

O processo de aquisição implica a necessidade de listagens de trabalho, do tipo lista de pedidos em andamento para cada tipo de material, lista de pedidos em andamento ordenada por número do pedido, lista de pedidos em andamento ordenada por fornecedor, carta de pedidos ao fornecedor, carta de reclamação ao fornecedor, carta acusando o recebimento do material, etc.

Caso o sistema seja totalmente on-line, parte destas listas (e outras) pode ser eliminada, podendo-se obter as informações diretamente no vídeo do terminal.

3.1.2. Emissão de catálogos

As informações geradas no processo de aquisição e complementadas com outras, resultado de atividades intelectuais como classificação e catalogação, permanecem gravadas num arquivo geral que pode ser utilizado por vários programas, para as finalidades específicas de cada um.

No caso da emissão de catálogos, haverá um ou vários programas cuja finalidade é acessar este arquivo, registro por registro (o que equivale a dizer documento por documento), retirando deles as informações necessárias para a preparação e a emissão dos catálogos previstos, bem como para a emissão de etiquetas para lombadas e para bolsos.

O que estes programas vão fazer é selecionar as informações necessárias, arranjá-las de acordo com o tipo de catálogo a ser emitido e efetuar a emissão conforme as especificações desejadas: em papel, em fichas, em microformas, e assim por diante.

O formato dos catálogos pode obedecer a normas internacionais ou não, dependendo das necessidades de intercâmbio que a biblioteca tenha.

A emissão de catálogos em microformas se dá, via de regra, quando a biblioteca pertence a uma rede ou a um sistema onde cada participante divulga seu acervo processado. Neste caso, a emissão pode ser efetuada diretamente pelo processo COM (computer output microform).

A emissão em fichas permite manter os catálogos sempre atualizados pela simples inserção de fichas novas nos arquivos existentes.

A emissão sob forma de listagens, embora a mais barata, tem o inconveniente da dificuldade de atualização.

De fato, se a cada nova entrada de material for efetuada uma emissão total, o processo se torna oneroso e demorado.

Uma alternativa é emitir suplementos dos catálogos, contendo somente o material recém-processado. Estes suplementos, relativamente pequenos, podem ser reimpressos com as atualizações do período e uma ou duas vezes ao ano pode ser feita a emissão global, incorporando os suplementos acumulados ao catálogo geral.

A vantagem da impressão de catálogos em listas é a possibilidade de divulgar o acervo da biblioteca dentro da empresa, da instituição, da escola, etc., se for o caso.

Os vários departamentos de uma instituição de pesquisa, por exemplo, podem receber catálogos impressos, tomando conhecimento do acervo da biblioteca, sem precisar se locomover até ela.

É bem verdade que existe sempre a possibilidade de se manter os catálogos da forma que for julgada mais conveniente e divulgar o acervo (novo ou global) por meio de um tipo de índice permutado (KWIC, KWOC, KWIT, etc.).

Testado o sistema e corrigidos os eventuais erros pode ser programada a implantação definitiva, a qual deve ocorrer na época mais oportuna e após todos os preparativos terem sido ultimados.

A implantação definitiva do novo sistema não implica a desativação do anterior. Pelo contrário, os dois devem seguir paralelamente por um tempo considerado suficiente para o novo sistema se tornar confiável. Só então o sistema anterior pode ser abandonado.

Para encerrar este item, queremos chamar a atenção para a necessidade de documentar cada passo, cada decisão, cada alternativa, cada *lay-out*, cada reunião, tudo enfim que tenha a ver com o sistema.

Esta documentação é importante por dois motivos: primeiro porque ela permite entender o sistema, seus porquês, a razão das decisões tomadas, dos recursos envolvidos, etc.; em segundo lugar, uma documentação bem-feita proporciona rapidez e segurança maiores, se e quando se fizerem necessárias alterações, tanto por parte da equipe que projetou o sistema como, com maior razão, por parte de outras pessoas que no futuro venham a ter de cuidar do mesmo. Estas pessoas só irão entender o sistema se ele estiver bem documentado.

A emissão destes índices pode ser efetuada a partir do mesmo arquivo geral, bastando que um programa, preparado para este fim, selecione os dados necessários e imprima o índice permutado.

Normalmente estes índices são utilizados para divulgar material novo. Nestes casos, o programa trabalharia os dados do material novo, antes que o mesmo fosse acrescentado ao arquivo geral.

Isto quer dizer que, nas situações onde é previsto um tipo de processamento específico para o material novo, este deve permanecer gravado separadamente, num arquivo provisório, até o processamento específico estar concluído e só então ser incorporado ao arquivo total.

3.1.3. *Circulação*

Sempre a partir do mesmo arquivo geral, mais um arquivo contendo dados dos consulentes da biblioteca, é possível efetuar um controle automático do material emprestado.

São fornecidos ao computador dados do usuário, do material e parâmetros que permitam calcular a data de devolução (admitindo que, dependendo do tipo de material, estes prazos possam ser diferentes).

Feito isto, o programa responsável pelo controle de empréstimos tem condições de calcular a data de devolução do material.

Chegando a data em questão, se a devolução do material não for acusada, o sistema pode emitir uma relação de usuários em atraso e/ou uma carta para cada um deles, solicitando a devolução.

Quando o material é devolvido no prazo, esta informação é passada ao computador o qual interrompe a "contagem regressiva".

Com este controle é possível saber, a qualquer momento, com quem está determinado material, quando vai ser devolvido e se há reservas para o mesmo.

A partir deste controle de empréstimos é possível obter, também automaticamente, informações do tipo:

- obras emprestadas num determinado período, relacionadas, por exemplo, em ordem decrescente de vezes em que foram emprestadas,
- relação de assuntos mais procurados para empréstimo,
- relação de usuários que levaram material emprestado,

- relação de usuários que nunca levaram material emprestado, num determinado período ou no total,
- relação de usuários que devolveram material com atraso, etc.

3.2. A DOCUMENTAÇÃO AUTOMATIZADA

A explosão documentária e a evolução dos sistemas de computador — tanto a nível de hardware como de software — fizeram com que fossem desenvolvidas novas formas, mais rápidas e eficazes, de captura, tratamento e difusão de informações.

Neste item trataremos de discutir alguns aspectos da automação das operações documentárias chamando a atenção para os pontos críticos e enfatizando mais as implicações da automação do que propriamente sua operacionalização.

3.2.1. *Indexação automatizada*

A indexação é, indiscutivelmente, uma das operações documentárias mais críticas e importantes.

Indexar um documento é determinar os assuntos principais que ele aborda e traduzir estes assuntos em forma de códigos, cabeçalhos de assunto ou palavras-chave, retirados de sistemas de classificação ou thesauri.

A utilização de computadores fez com que o uso de palavras-chave fosse preferido às demais formas de indexação.

Assim sendo um grande número de thesauri vêm sendo desenvolvidos buscando cobrir a maior parte das áreas do conhecimento humano.

A indexação automática dispensa a participação humana no processo.

Num primeiro momento, a idéia pode preocupar, na medida em que a indexação tem todas as características de uma atividade altamente intelectual, podendo ser satisfatoriamente desenvolvida apenas por especialistas na área tratada pelo documento a ser indexado.

Entretanto, a experiência tem mostrado que mesmo a indexação manual não garante que um documento venha a ser recuperado em resposta a uma questão formulada, ainda que ele seja pertinente; portanto, não se justifica o esforço e o alto custo de uma indexação manual.

Vale a pena, isto sim, dedicar os esforços no sentido de indexar as questões propostas pelos usuários, tendo em vista a forma como os documentos foram indexados pela máquina.

Existem dois tipos de indexação automática: com ou sem a utilização de thesauri.

A indexação automática sem a utilização de thesaurus é efetuada a partir de um antídicionário, ou seja, uma relação de palavras não significativas (artigos, preposições, advérbios, etc.).

Estas palavras são gravadas no computador e cada texto a ser indexado é comparado, palavra por palavra, com esta relação.

As palavras que não constarem da relação são consideradas possíveis palavras-chave ou descritores.

Esses descritores podem ser aceitos na totalidade ou submetidos a uma análise estatística fundamentada em critérios de frequência. Neste caso, é considerado descritor apenas a palavra que, além de não constar do antídicionário, atinja, pelo menos, uma frequência mínima.

A indexação automática efetuada com a utilização de um thesaurus implica a necessidade da preparação e gravação prévias deste thesaurus.

Os textos são comparados com o thesaurus gravado e as palavras com correspondência no thesaurus são consideradas descritores.

Cumprе salientar que os thesauri utilizados para indexação automática são mais extensos e mais detalhados do que aqueles usados na indexação manual.

Isto porque, na indexação manual, é possível reduzir as várias formas de um mesmo conceito a uma forma padronizada (por exemplo, computador, computadores, computação, computar, computacional, computacionais, e assim por diante), coisa que a máquina não é capaz de fazer. Assim sendo, o thesaurus a ser gravado possui as várias formas de um mesmo conceito registradas como sendo formas equivalentes. Este registro pode ser feito palavra por palavra ou utilizando raízes de palavras, desde que tomados certos cuidados

para não reunir num mesmo conceito palavras que nada têm a ver, apesar de terem mesma raiz (por exemplo, CONTR/ que pode reunir contrato, contratos, contratado, contratados, contratual, mas pode reunir também, controle, contrário, controlado, contrariamente, controvérsia, etc.)

A indexação automática, seja com ou sem uso de thesaurus, pode ser efetuada sobre um texto básico que pode ser um título, um resumo ou ainda um texto integral.

Efetuada sobre um título torna-se um processo econômico e rápido, porém de resultados discutíveis.

Estima-se que um bom título não contém mais de 30% dos conceitos abordados no documento completo.

Se for utilizado o resumo, que é a modalidade mais comum, o processo não é tão econômico, mas os resultados são bastante satisfatórios.

Calcula-se que um resumo contém entre 70% e 90% dos conceitos tratados no documento.

Quanto à utilização do texto integral, esta solução é tão onerosa que raramente é praticada.

Existem softwares de recuperação de informações que atuam sobre o texto como um todo, independentemente das palavras serem ou não descritores ou estarem reunidas num campo específico do registro gravado.

Para isto eles dispõem de uma vasta gama de operadores: lógicos, contextuais, relacionais e especiais (estes últimos para trabalhar com raízes ou terminações de palavras), que permitem equacionar uma questão a ser comparada com um texto em linguagem natural e não apenas com um conjunto de descritores ou, em outras palavras, uma linguagem documentária.

Se programas deste tipo forem empregados, desaparece, pelo menos aparentemente, a necessidade de indexar o texto.

Entretanto não é bem assim: em primeiro lugar, devemos considerar que o armazenamento de um resumo é mais demorado e caro do que o armazenamento de uma dezena de descritores. Em seguida, devemos considerar que é mais fácil e eficaz formular uma questão dirigida a um texto em linguagem documentária do que dirigida a um texto em linguagem natural, com os problemas de sinonímia, homonímia e outros, próprios deste tipo de linguagem.

Em contrapartida, um resumo, ao ser recuperado, é mais elucidativo e claro, quanto ao texto que ele resume, do que um conjunto de descritores o são quanto ao texto que eles indexam.

Caso seja decidido armazenar resumo e descritores, para um mesmo texto, como acontece em alguns casos, é interessante usar um programa de recuperação geral, que permita efetuar consultas em linguagem natural e em linguagem documentária, ao mesmo tempo ou isoladamente, conforme a necessidade.

Do que foi dito sobre indexação automatizada, como concluir o que é melhor?

Difícil dizer. Como sempre, o melhor é o que melhor atende nossas necessidades ou o que nos é possível fazer.

3.2.2. Recuperação automática de informações

Quando lidamos com recuperação de informações, devemos ter sempre em mente quatro aspectos fundamentais, de cujo equacionamento depende um maior ou menor grau de satisfação do usuário destas informações.

Temos os problemas de pertinência e de exaustividade, por um lado, e de rapidez e economia, por outro. Destes, os dois primeiros são problemas de ordem operacional e os outros dois, problemas de ordem organizacional.

Para os problemas de rapidez e economia é preciso que o sistema funcione de modo a permitir que o tempo existente entre o aparecimento da informação e sua inclusão numa coleção, passível de ser consultada, seja o menor possível. Além disso, o sistema deve fornecer a informação no menor tempo possível, quando esta vier a ser solicitada pelo usuário. Tudo isto, evidentemente, deve ser conseguido com um máximo de economia, não apenas para o usuário mas também para a coletividade à qual este pertence.

No que se refere à operacionalidade do sistema, temos os fatores pertinência e exaustividade que interferem diretamente no aspecto qualitativo da informação recuperada.

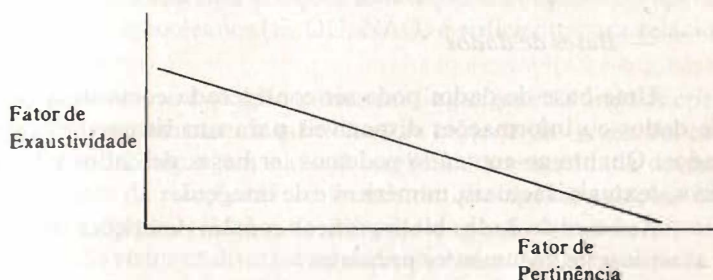
Existe um método, estabelecido por Cleverdon, para medir qualitativamente estes dois elementos.

	Informações obtidas	Informações não obtidas
Informações pertinentes	A	B
Informações não pertinentes	C	D

$$\text{Fator de pertinência} \dots\dots\dots \frac{A}{A + C}$$

$$\text{Fator de exaustividade} \dots\dots\dots \frac{A}{A + B}$$

Além disso Cleverdon estabeleceu uma lei segundo a qual a pertinência varia em função inversa da exaustividade.



Podemos conceituar pertinência como sendo a relação entre o número de informações pertinentes numa resposta e o total de informações desta resposta.

Exaustividade é a relação entre o número de informações pertinentes recuperadas e o total de informações pertinentes existente na coleção.

Para a obtenção de respostas com o maior fator de pertinência e de exaustividade possíveis é preciso conhecer basicamente:

- o assunto coberto pela coleção, de modo a poder avaliar a qualidade das informações recuperadas e, mais importante ainda, poder formular questões adequadamente;

— as características das informações e da coleção que as contém (estrutura, forma de indexação, etc.);

— todos os recursos do software de recuperação, utilizando-os da melhor forma possível.

A experiência tem mostrado que uma mesma estratégia, muito bem formulada tecnicamente, pode dar resultados diferentes quando aplicada em coleções distintas.

Ainda quanto à problemática envolvendo pertinência e exaustividade, podemos afirmar que uma questão aplicada a uma pequena coleção deve ser formulada enfatizando o aspecto exaustividade, preferindo recuperar algumas informações não pertinentes, mas tendo a certeza de que as informações pertinentes foram recuperadas totalmente.

Por outro lado, em se tratando de coleções volumosas, a ênfase deve ser dada ao aspecto pertinência, para evitar recuperar uma massa de informações tão grande que torne impossível seu aproveitamento.

— *Bases de dados*

Uma base de dados pode ser considerada como uma coleção de dados ou informações disponíveis para um sistema de computador. Quanto ao conteúdo podemos ter bases de dados bibliográficos, textuais, factuais, numéricos e de imagens.

As bases de dados bibliográficos contêm descrições sinaléticas e analíticas de documentos primários.

As textuais reúnem documentos primários propriamente ditos.

Bases de dados factuais possuem informações textuais não estruturadas como as textuais. Por exemplo, biografias, pesquisas, etc.

As bases de dados numéricos contêm dados como, por exemplo, preços, horários, estatísticas, etc.

Existem bases de dados contendo gráficos, desenhos, pinturas, enfim, imagens.

Atualmente existem no mundo milhares dessas bases, a maioria de dados bibliográficos, cobrindo praticamente todo o conhecimento humano.

No que se refere à distinção entre base e banco de dados, é

preciso saber que há autores para os quais não há diferenças, usando indistintamente um conceito pelo outro. Entretanto, há autores para os quais as bases de dados contêm descrições de documentos dentro dos quais é possível encontrar as informações desejadas, ao passo que bancos de dados contêm as próprias informações utilizáveis.

Para nós, um banco de dados é um conjunto de bases de dados, cada qual versando sobre determinado assunto. Assim sendo, o ORBIT, o DIALOG, o PASCAL, etc., são bancos de dados constituídos por várias bases de dados sobre física, música, religião, arte, história, etc.

— Estratégias de busca

Uma estratégia de busca é um conjunto de termos (descritores ou palavras da linguagem natural) relacionados entre si por meio de operadores.

Em se lidando com coleções indexadas o conjunto de operadores lógicos ou booleanos (E, OU, NÃO) é suficiente para relacionar os descritores.

Entretanto, se a coleção estiver em linguagem natural, os operadores lógicos se tornam insuficientes na maioria das vezes. As características da linguagem natural implicam a necessidade de um maior número de relações entre termos do que apenas relações lógicas. Por exemplo, relações de adjacência, de sinonímia e outras. Adiante serão vistos os diversos tipos de operadores existentes.

Seja em linguagem documentária ou em linguagem natural, toda questão devidamente preparada deve ser submetida ao computador, o qual, através de um software especial, analisa todos os documentos da base de dados, recuperando aqueles que satisfazem a questão submetida.



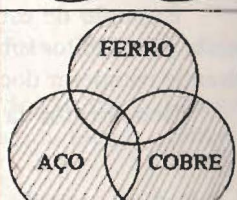
Exemplo de estratégia de busca: Numa base de dados contendo documentos sobre metais, são submetidas algumas estratégias visando recuperar documentos específicos.

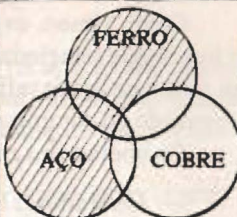
Documentos da base de dados:

A	B	C	D
Autor	Autor	Autor	Autor
Título	Título	Título	Título
Imprensa	Imprensa	Imprensa	Imprensa
Descritores	Descritores	Descritores	Descritores
..... Aço
..... Ouro	Cobre
Ferro

Vários usuários, com necessidades específicas, formulam suas perguntas na linguagem do dia-a-dia de cada um deles.

Estas questões são traduzidas em estratégias de busca e submetidas ao computador que recupera os documentos pertencentes aos universos definidos pela estratégia de busca.

Pergunta Formulada	Estratégia de busca	Universo recuperado
Desejaria recuperar documentos sobre ferro, cobre e aço	Ferro <i>E</i> aço <i>E</i> cobre	
Quero documentos que tratem de ferro e aço mas não de cobre	Ferro <i>E</i> aço <i>NÃO</i> cobre	
Pelo amor de Deus, preciso de documentos que tratem de ferro ou de aço ou cobre, qualquer coisa.	Ferro <i>OU</i> aço <i>OU</i> cobre	

Oh meu! Vê se tu tens uns bagulho que falem de ferro ou de aço, mas não me polui com cobre. Sacou bicho?	Ferro <i>OU</i> aço <i>NÃO</i> cobre	
--	--------------------------------------	---

Olhando a coluna “Universo recuperado” é fácil perceber que o operador E é restritivo, na medida em que ele obriga a presença simultânea dos descritores que ele relaciona.

Em contraposição, o operador OU fornece um conjunto de documentos-resposta bastante amplo, pois obriga a presença de pelo menos um dos descritores que relaciona.

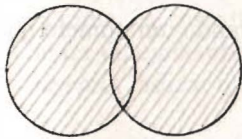
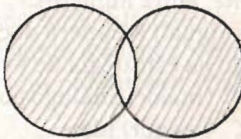
Com relação ao operador OU, se faz necessário um esclarecimento.

Existem softwares que adotam um tipo de operador OU, denominado OU exclusivo (representado por XOU, por exemplo).

Assim sendo, se relacionarmos os descritores ferro e cobre com o operador OU, recuperaremos documentos contendo o descritor ferro ou os documentos contendo o descritor cobre ou ainda os documentos contendo ambos descritores.

Entretanto se usarmos o operador XOU, recuperaremos os documentos contendo o descritor ferro ou os documentos contendo o descritor cobre, mas não os documentos contendo os dois descritores.

Esquematicamente, teríamos:

FERRO OU COBRE	FERRO XOU COBRE
	

— Operadores

Como vimos, os operadores são elementos de uma estratégia de busca que relacionam os termos desta estratégia de modo a obter um todo que seja o reflexo da necessidade do usuário e que possa ser submetido ao computador.

Os operadores mais freqüentes são:

- Operadores lógicos ou booleanos;
- Operadores contextuais;
- Operadores relacionais;
- Operadores especiais.

Operadores lógicos

Também conhecidos pelo nome de operadores booleanos por derivarem da álgebra de Boole, são os operadores:

E (intersecção)

OU (união) —XOU

NÃO (exclusão).

Operadores contextuais

São operadores mais utilizados na linguagem natural, que estabelecem relações de adjacência ou de proximidade entre palavras, ou relações de sinonímia.

Como exemplo, podemos citar os operadores:

ADJ — que obriga duas palavras a serem adjacentes;

SIN — que faz com que duas palavras sejam consideradas sinônimas na estratégia de busca onde ele é empregado.

Operadores relacionais

Estabelecem relações de grandeza entre palavras ou, mais precisamente, entre números.

= (igual) recupera documentos contendo um número igual ao fornecido na estratégia de busca.

> (maior) análogo ao anterior.

< (menor) idem.

Estes operadores são usados, normalmente, para restringir uma resposta obtida por uma consulta anterior.

Por exemplo:

1.^a estratégia: Silvio ADJ Santos (para recuperar documentos escritos por um autor com este nome. Esta estratégia estaria sendo submetida no campo de autoria.)

2.^a estratégia: 1 data/publ = 1925 (para selecionar dentre os documentos-resposta da estratégia 1, apenas aqueles cuja data de publicação é igual a 1925).

Operadores especiais

São, por exclusão, os operadores que não pertencem a nenhuma das categorias acima citadas.

Como exemplo, podemos citar o operador RAIZ que recupera documentos contendo palavras começando por uma raiz dada na estratégia.

Exemplo: RAIZ Bibliotec recuperará documentos contendo palavras como: biblioteca, biblioteconomia, bibliotecário, etc.

Neste tópico, denominado "operadores", foram abordados os tipos de operadores utilizados nas estratégias de busca e uns poucos exemplos de cada.

No anexo 4 teremos oportunidade de apresentar o STAIRS, que é um software para recuperação de informações, quando então serão vistos, detalhadamente, todos os operadores que ele possui, entendendo melhor o funcionamento de cada um deles.

— Parênteses

Numa estratégia de busca complexa é comum a utilização de parênteses para alterar a hierarquia existente entre os operadores, assim como acontece na matemática com as operações aritméticas.

Exemplo: $10 \times 2 + 3 = 23$ (primeiro efetua-se o produto e depois a soma, caso contrário o resultado seria diferente).

Se for desejado efetuar primeiramente a soma e depois o produto, devemos indicar o fato com o uso dos parênteses: $10(2 + 3) = 50$.

Nas estratégias de busca também é possível alterar a ordem de execução dos operadores usando os parênteses.

Normalmente a hierarquia dos operadores é a seguinte:

O operador E prevalece sobre o NÃO que, por sua vez, prevalece sobre o OU. (Isto falando somente dos operadores lógicos.)

Considerando a hierarquia geralmente adotada, uma estratégia de busca do tipo: “biblioteconomia OU documentação E informática”, faria com que o computador fornecesse, como satisfatórios, documentos que tratassem de biblioteconomia ou então documentos que tratassem de documentação e informática, ao mesmo tempo.

Caso desejássemos recuperar documentos que tratassem de biblioteconomia ou de documentação (indiferentemente), além de tratarem também de informática, então a estratégia deveria ser formulada da seguinte maneira: (biblioteconomia OU documentação) E informática.

— *Tipos de estratégias de busca*

Uma estratégia de busca pode ser submetida ao computador de forma ocasional ou periódica.

A forma ocasional, conhecida como busca retrospectiva, se dá sempre que um usuário necessita de uma determinada informação para atender uma necessidade imediata.

Então ele formula sua pergunta, a qual, após traduzida numa estratégia de busca, é submetida ao computador, que inicia a procura dos documentos pertinentes na base de dados.

O usuário recebe a resposta e o processo está encerrado até que nova necessidade surja e nova pergunta seja formulada.

A forma periódica, conhecida como Disseminação Seletiva da Informação — DSI (ou SDI, do inglês Selective Dissemination of Information) —, se dá de forma sistemática, sempre que documentos novos são incorporados à coleção.

O processo funciona da seguinte forma:

Cada usuário, dentro de sua área de interesse, formula uma ou mais perguntas que são traduzidas em estratégias de busca e gravadas em computador (após testada sua eficácia), com o código do usuário a que pertencem. É o que se chama de “Perfil do usuário”.

Toda vez que um material novo está para ser incorporado à coleção, ele será comparado com cada perfil e os documentos-resposta serão enviados ao usuário.

— Verificações

Todo processo automatizado requer certos cuidados e certas verificações constantes.

No caso de processamento automático da informação, os cuidados mais importantes são:

- Comparar os descritores atribuídos a cada documento com uma relação de descritores (thesaurus ou não), para evitar a entrada em computador de descritores errados, prejudicando as consultas.

- Verificar a ortografia do texto (título, resumo, etc.), para não gravar palavras erradas que nunca mais serão recuperadas (principalmente em se tratando de pesquisas em linguagem natural).

- Tomar cuidados no sentido de que cada tipo de informação (campo) ocupe o lugar certo dentro do documento (registro).

- Dispor de um meio automático de anotar o número de vezes que um documento foi recuperado, em resposta a questões formuladas, visando efetuar, periodicamente, um processamento que permita verificar quais documentos nunca (ou pouquíssimas vezes) foram recuperados e quais “sempre” são recuperados.

Uma análise destes documentos “raros” pode levar a conclusões interessantes como por exemplo: o documento está mal indexado (descritores atribuídos erroneamente ou indexação extremamente específica ou ainda indexação genérica demais); ou, então, o documento não pertence à área coberta pela coleção; e assim por diante.

De todas as verificações possíveis, das quais apenas demos alguns exemplos, a maior parte pode ser efetuada pelo próprio computador, através de programas especiais denominados genericamente de “programas de consistência”.

Outras, entretanto, por sua própria natureza, requerem intervenção humana.

4 ANEXOS

"O usuário é um mal necessário. Há que se ter paciência com ele, pois ele tem muita conosco."

4.1. ANEXO 1 — O ILS

Quando, por volta de 1968, o então Instituto de Energia Atômica considerou a necessidade de automatizar sua excelente biblioteca, foi decidido adotar o sistema utilizado na biblioteca do Centro de Processamento de Informação Científica da Comunidade Europeia, sediado em Ispira — Itália — o ILS — Integrated Library System.

Na época, este sistema era um dos raros de que se tinha notícia e, provavelmente, foi um dos primeiros sistemas integrados a serem implantados no Brasil.

Hoje, este sistema ainda existe naquela instituição de pesquisa, embora reformulado em alguns aspectos e acrescido de uma série de programas e rotinas, principalmente no que se refere à parte on-line, que foi totalmente desenvolvida lá.

Neste anexo será feita uma breve descrição do sistema, tal como veio de Ispira, para permitir ter uma idéia da forma como se pensava em automação de bibliotecas na década de 60.

Cumprе salientar que os dados para esta descrição foram retirados da publicação "O sistema de processamento de dados utilizado pela Divisão de Informação e Documentação Científicas do

Instituto de Pesquisas Energéticas e Nucleares", de Terezine Arantes Ferraz, Odete Guedes e Irene Josefa de Souza, datado de novembro de 1979.

Comentários eventuais do autor tentam ilustrar alguns aspectos por este considerados importantes, já que o mesmo teve oportunidade de participar dos estudos de viabilidade, bem como da fase inicial de implantação do referido sistema na época em que exerceu as funções de analista de sistemas no CPD daquela instituição.

O ILS foi concebido para automatizar as tarefas básicas de uma biblioteca, quais sejam: aquisição, processos técnicos, circulação e informação.

Por razões operacionais as publicações são agrupadas, de acordo com sua natureza, em "tipos" e assim serão tratadas pelo sistema e posteriormente representadas nos diferentes catálogos.

O sistema prevê 10 modalidades distintas de aquisição, a saber:

SING	Pedido avulso
SUBS	Assinatura
STOR	Pedido permanente
SECT	Pedido por secções
COMP	Pedido global
MEMB	Anuidades
FREE	Grátis
EXCH	Permuta
LOAN	Circulação
BACK	Material já existente na biblioteca, a ser incorporado ao sistema (backlog).

O ILS prevê um conjunto de relatórios divididos em 4 grandes grupos:

1º) Relatórios para fins de aquisição

Lista on-order, para monografias, séries monográficas, teses, etc.

Lista on-order, para periódicos.

Lista on-order, para séries periódicas.

Lista on-order, para relatórios.

Lista on-order, classificada por número de pedido.

Lista on-order, classificada por nome de fornecedor.

Carta de solicitação para fornecedores.

Carta de reclamação para fornecedores.

Carta acusando recebimento do material.

Lista de periódicos, classificada pelo título/fornecedor.

Lista de periódicos, classificada pelo fornecedor/título.

Lista das publicações em reserva, classificadas pelo número de chamada.

Lista por título de periódicos conservados num determinado departamento da instituição. (Era adotado o esquema de empréstimo permanente para determinados departamentos da instituição, para os quais o material era adquirido.)

2º) Relatórios para fins contábeis

Lista de liquidação de faturas, classificada por fornecedor/número de pedido.

Lista de publicações ainda não pagas, mas já recebidas.

3º) Relatórios para procedimentos de empréstimos

Lista de todas as publicações emprestadas, classificada pelo número de chamada do livro.

Lista de publicações em empréstimo permanente, classificada por departamento e/ou usuário.

Lista de reclamações por usuário.

4º) Relatórios para fins estatísticos (vários)

Os catálogos previstos pelo ILS são:

1) Catálogo principal. Classificado numa única ordem alfabética de autor e título. Este catálogo contém as seguintes indicações: autor, título, imprensa, colação e número de chamada.

2) Catálogo de assunto. Classificado alfabeticamente de acordo com as 61 grandes categorias de assunto constantes do INIS Chart (1ª hierarquia); dentro destas grandes categorias, ainda em ordem alfabética, aparecem os descritores extraídos do Thesaurus INIS (2ª e 3ª hierarquias). Dentro destas hierarquias os livros são classificados em ordem decrescente de data. (Observação: INIS é a sigla de International Nuclear Information System.)

3) Índice de assuntos. Classificado alfabeticamente por descritores.

4) Catálogo de conferências. Este catálogo se subdivide em 3 partes: Título: por título do evento. Data: por data do evento e Local: por local de realização do evento.

5) Catálogo de séries monográficas. Classificado por título da série e dentro desta em ordem crescente do número dos volumes.

6) Catálogo de autores corporativos. Classificado por nome do autor corporativo e dentro deste por título da obra.

7) Catálogo de trabalho. Classificado de acordo com a Classificação Decimal Universal — CDU.

8) Catálogo de publicações recém-adquiridas.

9) Catálogo de periódicos. Contendo todos os dados bibliográficos, incluindo transcrição da coleção.

10) Lista especial de informações sobre determinados assuntos. Esta lista, baseada nos perfis dos usuários, divulga informações sobre publicações recém-chegadas.

Devido à natureza variada das publicações e conseqüente variedade de processamento, as publicações foram divididas em:

Monografias

Séries monográficas

Periódicos

Séries periódicas

Relatórios

Opus

Teses

Panfletos

Analíticas.

Para que o produto final corresponda às exigências da biblioteca, as informações bibliográficas são fornecidas ao sistema distribuídas em 9 categorias:

Identificação

Descritores

Autoria

Título

Imprensa

Outros dados bibliográficos

Número de chamada (CDU mais Cutter)

Vaga

Dados administrativos.

O problema da variabilidade do tamanho dos dados, como também da própria presença dos mesmos, foi deixado a cargo do programa, fazendo com que a transcrição para a planilha de entrada fosse simples e prática e o trabalho de perfuração limitado ao essencial.

Os campos pertencentes a cada categoria são separados por ponto e vírgula. A ausência de determinado campo é acusada pela presença do separador e o tamanho de cada campo é limitado aos caracteres significativos presentes. O alinhamento e a expansão dos campos em formatos fixos são executados pelo programa por meio de tabelas que especificam as características dos campos dentro de cada categoria.

Portanto, as categorias previstas para a entrada de dados e a indicação das respectivas funções são as seguintes:

1	Identificação para ordens ou modificação
1A	Identificação para recebimento
1AU	Identificação para recebimento de adendos de atualização
1B	Informação bibliográfica
1C	Identificação para notas de reclamação ao fornecedor
11L	Ligação para séries não numeradas
11R	Referências cruzadas
12	Número de localização de biblioteca dependente (previsto no sistema original, mas não utilizado aqui)
13	Número de referência externa: relatórios, patentes
2	Classificação por assunto
22	Classificação de biblioteca dependente (não usado)
2KL	Palavras-chave
3N	Autor(es) nominativo(s)
3C	Autor(es) coletivo(s)
4L	Título
4LS	Subtítulo
4LC	Subtítulo
4LN	Novo título (mudança de título, para periódicos)
5I	Imprensa
5II	Paginação

6B	Dados bibliográficos
6C	Dados de conferência
6H	Acervo de periódicos e séries periódicas
7	Número de localização
91	Ordem
911	Número de ordem para ser referida para ordem compacta
92	Previsão para data de recebimento
931/4	Reserva
935	Circulação
94	Notificação para "tabela de conteúdo"
951/4	Empréstimo
961/4	Devolução
9A	Recebimento de publicação (informações suplementares)
9I	Data de fatura
AD1/3	Descrição do fornecedor
ADP	Descrição do departamento/serviço (não usado)
AU	Descrição do usuário
AS/P	Descrição do assunto
ARN/E	Referências gerais
ASN/E	Descritores

Os dados bibliográficos são introduzidos no sistema por ocasião da aquisição da primeira cópia, mediante o preenchimento da planilha de entrada.

Com os elementos assim introduzidos, obtém-se os dados para efetuar o pedido ao fornecedor. Esses dados formam, também, a base para completar, posteriormente, a catalogação.

Ao chegar a publicação, nova planilha de entrada, alterando ou acrescentando dados novos àqueles fornecidos no momento do pedido. É nessa ocasião, com o material em mãos, que se completa a catalogação, sendo, então, o livro classificado pela CDU e indexado pelo thesaurus INIS.

Os dados bibliográficos são gravados num arquivo central, único, denominado "Master File", do qual são extraídos todos os dados necessários à execução das várias operações concebidas pelo sistema.

Por ocasião da gravação dos dados, os mesmos são submetidos a um programa de consistência, o qual emite mensagens de erro

quando ocorrem divergências entre as regras do ILS e o preenchimento das planilhas de entrada.

Semanalmente são emitidos catálogos parciais e na última semana do mês há uma emissão cumulativa.

Juntamente com os catálogos semanais, o sistema emite etiquetas correspondentes ao material relacionado nestes catálogos.

No final do ano, após uma série de verificações e eventuais alterações, é efetuada a emissão dos catálogos globais, cobrindo todo o período e incorporando os dos anos anteriores.

O empréstimo é efetuado utilizando um terminal de computador, instalado no balcão de empréstimos, que permite o registro on-line de todas as transações referentes a este processo.

Além disso, o programa de empréstimos prevê a emissão de estatísticas diárias do movimento de empréstimos, bem como a emissão de reclamações a usuários com material cujo prazo de devolução esteja vencido, etc.

No ato do empréstimo, o encarregado digita as informações, confere os dados na tela e carimba, no impresso colado à capa interna do livro, a data de devolução. O programa informa se o livro desejado está disponível ou não, se há reservas para o mesmo, se o usuário está dentro do limite de unidades às quais tem direito, se está em dia com as devoluções, etc.

O ILS é estruturado em 4 grandes módulos a saber:

- Módulo de tratamento de entrada, com cerca de 21 programas.
- Módulo de elaboração de catálogos, também com 21 programas.
- Módulo de controle de empréstimos, com 31 programas.
- Módulo de controle de assinaturas, com 4 programas.

Isto nos permite verificar a complexidade de um sistema integrado de automação de bibliotecas, já que aproximadamente 80 programas estão envolvidos.

Este sistema, embora de concepção em certos aspectos superada, está funcionando satisfatoriamente há anos, o que nos permite, mais uma vez, afirmar que "sistema bom é aquele que atende nossas necessidades". Além disso, é preciso ter em mente que em processamento eletrônico de dados a obsolescência é fato comum. Por outro lado, se ficarmos esperando sempre o aparecimento de hard-

wares e/ou softwares melhores, mais sofisticados e mais modernos, provavelmente nunca faremos nada.

4.2. ANEXO 2 — O DOBIS

O DOBIS é um sistema de automação de bibliotecas moderno e completo constituído de quatro módulos básicos:

Dortmund Library System,
Leuven Library System — On-line,
Leuven Library System — Batch,
Periodical Control System.

O DOBIS é designado para bibliotecas públicas, universitárias, especializadas e nacionais, unitárias ou em rede.

Ele integra as principais funções (serviços) de uma biblioteca.

Seus módulos on-line cobrem as tarefas de aquisição, catalogação, circulação, controle de periódicos e busca. Isto é, o DOBIS permite acessar o acervo de uma biblioteca, submetendo ao mesmo estratégias de busca.

Além desta parte on-line existem programas que permitem obter fichas catalográficas, saídas pelo sistema COM (Computer Output Microform), relatórios e listagens estatísticas e administrativas.

Com a utilização do DOBIS é possível:

- aumentar a produtividade do pessoal da biblioteca por tratar-se de uma ferramenta fácil de se aprender e de se utilizar;

- reduzir o tempo da catalogação por eliminar duplicação de trabalho e integrando fichário de autoridade;

- reduzir o tempo requerido para a circulação e aquisição de periódicos, pela eliminação do processo de datilografia e de arquivamento, acelerando a entrada das informações, provendo telas formatadas;

- obter listagens necessárias para usuários dentro e fora da rede.

A seguir, serão apresentadas de forma esquemática as características principais do sistema DOBIS, com a finalidade de dar uma idéia de sua potencialidade e recursos.

1. Aquisição on-line

— Armazenamento de dados de interesse para a aquisição. Caso o documento venha a ser adquirido os dados são transferidos para a catalogação e demais serviços.

— A posição do documento adquirido é indexada por vendedor, verba, nota fiscal, número do documento e índices bibliográficos.

— Os valores em moedas estrangeiras são convertidos automaticamente, eliminando cálculos manuais.

— As assinaturas que necessitam de renovação podem ser listadas. O próprio sistema prepara as renovações.

— O sistema credita e debita notas e pagamentos a vendedores e ajusta tudo às verbas da biblioteca, colocando à disposição estatísticas sobre vendas e verbas.

2. Catalogação on-line

— Registros catalográficos podem ser inseridos e modificados on-line e imediatamente (Real Time).

— O formato MARC (MACHine Readable Cataloging) é o suporte de monografias, seriados, mapas, manuscritos e música.

— Com o objetivo de melhorar a qualidade da catalogação, os arquivos de autoridade possuem referências cruzadas e notas, remetendo termos procurados aos termos adotados.

3. Circulação on-line

— O controle de empréstimos pode ser efetuado via terminal, existindo, para isso, um conjunto de parâmetros a serem utilizados segundo a política de empréstimo de cada biblioteca.

— Datas de devolução, atrasos, etc., são calculados automaticamente.

— Através de uma simples busca é possível obter a posição de um documento em circulação, quer pelo seu título, quer pelo nome ou número de inscrição do consulente.

4. Busca on-line

— Os registros bibliográficos são indexados por:

Autor
Título
Assunto
Número de classificação
Palavras significativas do resumo
Editor
ISBN
ISSN
Número de registro na bibliografia do país.

— Pode ser efetuada a pesquisa num catálogo com os 9 índices acima descritos, formulando questões com o emprego dos operadores lógicos ou booleanos: AND, OR e NOT (E, OU e NÃO respectivamente).

— Permite truncamento de termos minimizando a necessidade de chaves exatas para recuperar determinadas informações.

— Usuários autorizados podem acessar registros no que diz respeito à circulação, aquisição, através de itens como: indivíduo que tomou emprestado o documento, verba da biblioteca, editor, etc.

5. Acesso por parte do público, on-line

— Telas explicativas simplificam a busca do consulente.

— A estratégia de busca pode ser alterada durante a consulta.

— Uma tela explicativa fornece a posição do documento.

Quem o possui, condições de disponibilidade para empréstimo, etc.

— Requisições de usuários sobre vários documentos podem ser impressas imediatamente.

— Os consulentes podem enviar mensagens aos funcionários da biblioteca.

— Os consulentes podem “chamar”, na tela, partes substanciais do documento.

6. Resumos

— Resumos, em diversas línguas, podem ser armazenados e recuperados on-line.

— Termos de resumos são extraídos on-line, podendo ser pesquisados.

7. *Produção de catálogos*

— Programas "batch" produzem catálogos impressos ou em microformas (pelo sistema COM). Os catálogos impressos podem ser listados em fichas ou em papel normal de computador.

8. *Vários*

— São possíveis vários tipos de estatísticas do tipo: volume de empréstimo num determinado período, tipo de documentos que circularam num período, etc.

— O formato e o conteúdo dos relatórios e listagens são especificados on-line pelas bibliotecas, de forma independente, ainda que as mesmas pertençam a uma rede.

— As fitas magnéticas MARC podem ser amplamente utilizadas graças a uma interface padronizada.

4.3. ANEXO 3 — O VIDEOTEXTO

Conceitualmente, o videotexto é um serviço interativo, pelo qual os usuários podem acessar um banco de informações utilizando seu próprio telefone, seu receptor de televisão e um adaptador especial.

Quem primeiro viabilizou o videotexto foi a Inglaterra, por volta dos anos 70, onde recebeu, inicialmente, o nome de Viewdata, sendo conhecido hoje por Prestel.

Atualmente existem 4 tecnologias diferentes de Videotexto: Prestel, da Inglaterra; Antiope, da França; Telidon, do Canadá e o Captain, do Japão.

O Brasil adotou o sistema Antiope francês, por sua qualidade e custo.

Independentemente da tecnologia adotada, o videotexto apresenta a grande vantagem de poder funcionar com equipamentos que, via de regra, a população urbana já possui. De fato, é muito difícil encontrar alguém nas grandes cidades que não possua um telefone e um aparelho de TV.

Assim sendo, as linhas telefônicas são aproveitadas para a transmissão de informações que são solicitadas por meio de um adaptador especial e exibidas na tela da TV.

Outra grande vantagem é a facilidade de uso, já que nenhum conhecimento prévio é exigido, assim como não é necessário qualquer curso de operação.

O videotexto é auto-explicativo e ao ser chamado exibe na tela da TV uma relação completa das informações disponíveis no dia, associadas às teclas do adaptador especial.

Desta forma, se o usuário desejar informações sobre esporte, pressiona a tecla associada ao item esporte e a tela é substituída por outra contendo um índice específico ao assunto esporte, com itens do tipo: futebol, tênis, basquete, turfe, etc., onde a cada item também está associada uma tecla para que o usuário faça a opção desejada.

Chegando na informação desejada, o usuário pode efetuar a leitura tranqüilamente, mudando de tela (o que equivale a dizer "virando a página") pela simples pressão de uma tecla.

Trata-se, portanto, de um informativo eletrônico, com possibilidade de gráficos e figuras a cores e com uma disponibilidade de informações diferentes muito grande.

O custo do serviço é calculado automaticamente e o total debitado juntamente com a conta telefônica normal.

As informações disponíveis neste "jornal eletrônico" vão desde a consulta a um saldo bancário até roteiros turísticos, passando por informações meteorológicas, cotações da bolsa, filmes em cartaz e congressos.

O funcionamento do videotexto está baseado no trinômio: operador, fornecedor e usuário.

O operador é quem monta e administra o sistema, colocando à disposição dos fornecedores terminais de edição e o computador central e suprindo os usuários dos adaptadores necessários.

Os fornecedores são os responsáveis pela criação dos serviços que serão colocados à disposição dos usuários. Cabe a eles suprir o sistema com informações processadas dentro de uma determinada estrutura.

Os usuários são os que utilizam o serviço por meio de terminais domésticos.

Apesar do videotexto ser um serviço eminentemente público, é possível a formação de sistemas de videotexto veiculando informações destinadas a certos segmentos específicos da sociedade como,

por exemplo, hospitais, companhias aéreas, agências de turismo, sindicatos, bibliotecas, etc.

Que uso de videotexto poderia ser feito por uma biblioteca?

Em primeiro lugar uma biblioteca pode usar o videotexto como um usuário qualquer, colocando à disposição de seus leitores as informações que normalmente são veiculadas por este sistema.

Em segundo lugar a biblioteca, ou conjunto de bibliotecas, pode se tornar um fornecedor de informações, oferecendo aos leitores e aos funcionários informações do tipo: relação de aquisições recentes, catálogos, normas da biblioteca, títulos de periódicos assinados, cadastro de fornecedores, cadastro de usuários, relação de obras em processo de aquisição, relação de usuários em atraso, relação de bibliotecas congêneres, material solicitado e não constante do acervo, estatísticas diversas, calendário de eventos internos, calendário de eventos externos na área de atuação da biblioteca, resumos, etc.

Quer como usuário, quer como fornecedor, a biblioteca pode fazer do videotexto uma opção, uma alternativa ou mesmo uma complementação de seus serviços.

4.4. ANEXO 4 — O ENDS

O ENDS (European Nuclear Documentation System) é, como o próprio nome diz, um sistema de documentação nuclear, desenvolvido pelo CID (Centre of Information and Documentation), da EURATOM.

A referência a este sistema se deve, mais do que outra coisa, ao seu pioneirismo.

De fato, o início de suas atividades automatizadas data de 1964, sendo que desde 1961 o sistema já vinha funcionando de forma manual.

Em 1964, com a aquisição de um computador IBM-1401, o ENDS se automatizou e passou a processar milhões de informações no campo nuclear e afins.

Em 1966, o IBM-1401 foi substituído por um possante (na época) IBM/360 modelo 30 e poucos meses depois por um modelo 40 que dispunha de 12 unidades de fitas magnéticas, 2 unidades de

discos magnéticos e unidades impressoras e leitoras/perfuradoras de cartões.

O ENDS e os especialistas que nele trabalharam foram um marco na documentação automatizada e raros são os sistemas de documentação modernos que não adotem seus princípios ou que nele não tenham se inspirado. Difícil encontrar trabalhos sobre thesaurus que não falem da "Terminology Chart" usada no ENDS. Ainda hoje, nas mesas-redondas, reuniões e congressos sobre documentação automatizada, é obrigatória a presença de nomes que atuaram neste sistema. É por isso que consideramos oportuno falar dele neste anexo.

Vamos apresentar o ENDS sob forma de tópicos, para perceber melhor a sequência das operações.

— *Seleção das publicações secundárias*

A biblioteca, que não pertence ao ENDS, adquire, recebe e fornece ao sistema uma média de 60 títulos de fontes secundárias sobre energia nuclear e áreas correlatas.

Os critérios adotados pelo ENDS para selecionar estas publicações são, basicamente: a quantidade de resumos de interesse para o sistema, existentes na publicação; a qualidade destes resumos e o tempo empregado para divulgá-los.

Outro critério importante é a língua em que a fonte secundária é publicada. Observadas sob este enfoque, as publicações secundárias de maior prioridade são as que se apresentam em língua inglesa, seguidas daquelas em francês e finalmente as de língua alemã.

— *Registro em kardex*

Assim que as publicações secundárias são recebidas, o ENDS inicia uma série de controles, efetuados em kardex manuais.

Cada publicação possui uma ficha contendo os diferentes controles executados desde o momento da chegada até a colocação dos dados no computador.

A grande importância deste kardex é que ele permite:

— determinar onde se encontra determinada publicação e, conseqüentemente, por quais processos já passou e por quais ainda não;

- obter dados estatísticos;
- controlar a eficácia, em termos de tempo, do pessoal adido a cada tarefa.

— *Seleção dos resumos*

Uma vez efetuado o controle das publicações, estas são enviadas a especialistas que selecionam quais resumos têm interesse para o sistema.

Determinadas publicações possuem apenas capítulos específicos onde podem ser encontrados resumos de interesse. Em compensação, existem publicações que devem ser examinadas integralmente, pois, potencialmente, todo conteúdo interessa.

Os especialistas que efetuam a seleção não pertencem ao corpo de funcionários do sistema e executam o trabalho como free-lancers.

— *Codificação*

Quando o ENDS recebe de volta as publicações com os resumos de interesse assinalados, ele as envia a outra equipe externa para efetuar a codificação de cada resumo.

Esta codificação consiste em considerar certos elementos bibliográficos que caracterizam o documento primário, como: nome do autor, ano de publicação, paginação, e manipulá-los segundo normas preestabelecidas, de modo que cada resumo receba um mesmo código, independentemente da publicação secundária em que aparece.

Este código permite ao computador efetuar um controle de duplicatas entre o material novo e o que já se encontra na coleção.

Se o codificador tiver dúvidas quanto ao código a ser atribuído a determinado resumo, as instruções são para atribuir, ao mesmo, tantos códigos quantos sejam necessários para eliminar toda e qualquer dúvida.

— *Controle de duplicatas*

O material codificado é transcrito, em fitas magnéticas, para entrar no computador onde um programa específico executa o controle de duplicatas.

Este controle se dá em dois níveis; um deles se refere a uma comparação entre o material novo e o material existente; outro con-

fronto se dá a nível do próprio material que está por entrar no sistema, já que neste também pode haver duplicatas.

A exclusão das duplicatas é feita segundo um critério de prioridade baseado na qualidade do resumo e na língua em que o mesmo se apresenta.

— *Indexação*

Os resumos dados como novos pelo controle de duplicatas são enviados a especialistas (professores e pesquisadores) para indexar seu conteúdo.

Esta indexação é efetuada com o auxílio de um thesaurus que encerra toda a terminologia nuclear e afim, incluindo termos considerados proibidos, os quais remetem a termos adotados.

O indexador anota as palavras-chave numa ficha que contém os dados do resumo ao qual a indexação se refere.

Ê facultado ao indexador propor palavras-chave novas, as quais serão estudadas pela equipe do ENDS que cuida da atualização do thesaurus e, se aprovadas, passarão a constar das novas versões deste.

O resultado do trabalho de indexação é enviado ao ENDS que efetua controles do tipo: verificar se as palavras pertencem à terminologia nuclear. Caso contrário, é feita uma verificação para determinar se o resumo não pertence ao campo nuclear ou se a indexação foi malfeita. Outra verificação diz respeito ao número de palavras-chave atribuídas a cada resumo. Se este número for muito pequeno (menor que 5) ou muito grande (maior que 25) será feita uma verificação para estabelecer se o mesmo se justifica ou não.

Terminados estes controles e efetuados os eventuais acertos, o material é colocado no computador que irá efetuar um controle de terminologia por meio de um sistema denominado PIT (Processing Index Terms), que consiste em verificar se as palavras-chave atribuídas estão contidas no thesaurus com o qual são comparadas.

Certos termos, parcialmente errados, são corrigidos automaticamente, tomadas certas precauções.

Por exemplo, a palavra-chave URANIUM pode ser corrigida automaticamente se aparecer como:

URAIUM (omissão de uma letra),

URLANIUM (introdução de uma letra parasita),

UBANIUM (substituição de uma letra por outra),
URANUIM (inversão de letras consecutivas).

Em caso de dúvidas, o programa não efetuará a correção, listando a palavra-chave errada para que seja efetuada uma verificação e uma correção manuais. É o caso do termo errado RANIUM que não será corrigido por poder ser tanto a forma errada de URANIUM como de RADIUM.

— *Inclusão na coleção*

Assim que tudo estiver em ordem, a referência bibliográfica e as palavras-chave são introduzidas no computador.

O resumo correspondente é colado em fichas-padrão, microfilmado e armazenado em dispositivos denominados "CARD READER/PRINTER" da IMAGE SYSTEM, conectados ao computador de modo que, ao formular uma questão, o computador efetua a busca em base às palavras-chave, e o que for julgado pertinente é localizado, por meio de um código de amarração, e exibido na tela do equipamento, aumentado até 24 vezes.

— *Normalização da terminologia*

A normalização da terminologia foi considerada necessária pelo CID, apesar dos estudos que ela implicava.

Esta normalização adotada, não apenas para facilitar as consultas como também para assegurar um melhor controle de sua própria evolução, acabaria resultando num thesaurus contendo mais de 16.000 termos em suas edições mais recentes.

A estrutura interna do thesaurus divide os termos em duas categorias: termos positivos, a serem utilizados pelos indexadores no momento da análise do documento, e termos negativos (antecedidos por um traço no thesaurus), proibidos para indexação.

Os termos negativos se repartem em termos de operador USE e termos de operador VEJA.

Os primeiros cuidam dos casos de sinonímia, quase sinonímia e vizinhança semântica. Neste caso, o termo não é utilizado, adotando em seu lugar o(s) termo(s) que segue(m) o operador USE.

Exemplo: COLUMBIUM USE NIOBIUM.

Os segundos resolvem problemas de termos com significado

variável dependendo do contexto onde estão inseridos. Assim o termo deve ser substituído por um dos termos que seguem o operador VEJA, conforme o contexto.

Exemplo: COLUMNS VEJA CHROMATOGRAPHY

ou EXTRACTION COLUMNS

ou ION EXCHANGE MATERIAL

ou MECHANICAL STRUCTURES

ou etc.

Visando facilitar o trabalho do indexador foi elaborada uma estrutura gráfica na qual os termos foram repartidos em 57 quadros ou diagramas terminológicos, conhecidos como "Terminology Charts".

— Recuperação das informações

Esta recuperação, simples em princípio, significa traduzir a questão posta pelo usuário em termos de indexação e submetê-la ao computador para que ele selecione todos os documentos nos quais o conjunto de descritores inclua aqueles da questão assim traduzida.

Esta tradução pode se dar em dois níveis: a pesquisa de todas as possibilidades e alternativas de termos de indexação, de modo a obter uma resposta tão exaustiva quanto possível, ou a combinação destes termos por meio dos operadores lógicos: AND (X), OR (+), NOT (—).

Para melhor entendimento do mecanismo de formulação de questões, vamos recorrer a um exemplo que a própria literatura sobre o ENDS utiliza e que pedimos ao autor, M. Marcel Detant, permissão para reproduzir.

Trata-se de uma questão aparentemente simples que concerne à contaminação de animais pelo plutônio. Nesta questão encontramos três elementos: o objeto (animais), a causa (plutônio) e o efeito (contaminação), sobre os quais a atenção do indexador irá se fixar.

No que se refere ao objeto, não é suficiente introduzir a palavra-chave "ANIMALS" na questão. Animais específicos também devem constar da formulação, como por exemplo: DOGS, MICE, RATS, etc.

No diagrama terminológico 03 "ANIMALS" podem ser encontradas as palavras-chave apropriadas. Uma das saídas deste dia-

grama, "EMBRYOS", conduzirá ao diagrama 08 "GENETICS", onde pode ser considerado importante introduzir na questão EMBRYOS, FETUSES, GONADS, todos elementos passíveis de sentirem os efeitos de uma contaminação. Uma das saídas do diagrama 08 "GLANDS" remeterá ao diagrama 04 "ANATOMY", onde podem ser encontradas palavras-chave interessantes, tais como KIDNEYS, LIVER, LUNGS, etc.

Este exemplo nos permite apreciar a grande utilidade dos diagramas terminológicos que permitem desenvolver ao máximo o assunto a ser tratado e os vários aspectos do mesmo.

Podemos concluir que uma parte da questão a ser submetida ao computador se apresentará sob a forma:

$A1 + A2 + A3 + \dots + An$, onde: $A1 = \text{ANIMALS}$

$A2 = \text{DOGS}$

$A3 = \text{MICE}$

•
•
•

A causa (plutônio) será tratada da mesma forma, ainda que a situação seja mais simples no seu conjunto, na medida em que será suficiente pensar no plutônio e seus compostos, seus complexos e seus isótopos, o que constitui uma lista de palavras-chave relativamente importantes.

Esta parte da questão se apresentará, igualmente, sob a forma:

$B1 + B2 + B3 + \dots + Bn$, onde: $B1 = \text{PLUTONIUM}$

$B2 = \text{PLUTONIUM}$
 COMPOUNDS

•
•
•

O efeito (contaminação) não será tão detalhado ainda que sua interpretação em termos de palavras-chave seja bastante delicada, considerando que no metabolismo intervêm termos tais como: RETENTION, EXCRETION, METABOLISM, etc.

Aqui a formulação também será:

$$C1 + C2 + C3 + \dots + Cn, \text{ onde: } C1 = \text{CONTAMINA-} \\ \text{TION} \\ C2 = \text{METABOLISM}$$

* *

: :

. .

Finalmente, a questão global se apresentará na forma:

$$(A1 + A2 + A3 + \dots) \times (B1 + E2 + B3 + \dots) \times \\ (C1 + C2 + C3 + \dots),$$

ou seja, três grupos de alternativas relacionadas pelo operador AND (X).

A operação de tradução tornar-se-ia cansativa e morosa para o documentalista, se este tivesse que transcrever longas listas de animais, de compostos ou de isótopos.

Por esta razão foi prevista uma forma de agrupar conjuntos de palavras-chave de uma mesma família, atribuindo termos ainda mais genéricos, quais sejam, os geradores "ALL".

Desta forma, a questão global acima pode ser apresentada usando os geradores: ALL ANIMALS, ALL PLUTONIUM COMPOUNDS, ALL PLUTONIUM ISOTOPES...

Com estes geradores é possível formular uma questão visando recuperar documentos onde apareça a palavra-chave "animais" simplesmente e outra visando recuperar documentos contendo a palavra-chave "animais" bem como os documentos contendo os nomes dos próprios animais como palavras-chave.

No primeiro caso bastaria transcrever: ANIMALS e no segundo seria preciso usar ALL ANIMALS.

A recuperação das informações no ENDS é efetuada de duas formas: a busca retrospectiva, eventual e cobrindo a coleção completa, e a disseminação seletiva da informação (DSI), periódica (mensal) e usando apenas o material adquirido no período.

— *Performances do sistema*

1) O fator de pertinência é calculado para cada questão submetida, após a avaliação do documentalista, baseado nos resultados fornecidos pelo computador.

Não é fácil falar de um fator de pertinência médio do sistema. Dependendo das circunstâncias, o documentalista será levado a formular questões amplas que acarretarão, conseqüentemente, um fator de pertinência baixo. Há que se considerar também que as respostas às questões de DSI, possuem um fator de pertinência mais elevado que aquelas das buscas retrospectivas.

Assim sendo, o ENDS calcula que o fator de pertinência para as buscas retrospectivas será em média de 35%, ao passo que o do DSI se situa pelos 75%.

2) O fator de exaustividade é difícil de calcular pois, sendo ele a relação entre documentos pertinentes encontrados numa resposta e os documentos pertinentes existentes na coleção, seria preciso conhecer primeiro este último, para poder calcular o fator. A única forma de obter o fator de exaustividade exato seria a forma manual, impossível de realizar numa coleção com mais de um milhão de unidades. Entretanto, utilizando um método proposto por M. Rolling, baseado numa curva fator de exaustividade—fator de pertinência, de Cleverdon, o ENDS consegue determinar o fator de exaustividade, que se situa por volta de 75% quando o fator de pertinência é de 65%.

3) O fator economia é constituído essencialmente dos custos operacionais. Deixando os valores absolutos, pois, além de serem em francos belgas, datam de 1969, vale a pena mostrar a relação entre os custos de entrada de um documento no sistema e os custos para sua recuperação. Assim, estes últimos chegam a ser 50 vezes maiores que os primeiros.

4) O fator demora deve ser considerado tanto do ponto de vista da alimentação como da recuperação.

Na alimentação há uma demora inevitável e independente do sistema, conseqüência do tempo entre a aparição de um documento primário e sua referência na literatura secundária (que é sobre a qual o ENDS trabalha). Este tempo varia muito dependendo do periódico secundário considerado. Depois que um documento primário é referenciado numa publicação secundária, ainda existe o tempo de seleção, codificação, perfuração, controle de duplicação, indexação, processo PIT. Destes, a seleção, a codificação, a perfuração e a indexação são efetuados fora, por pessoas contratadas. Isto aumenta o tempo de introdução no sistema que chega a 3 meses aproximadamente.

O tempo de recuperação inclui: tempo de formulação, tempo de computador, tempo de busca dos resumos nas microfichas, tempo de avaliação e de fotocópia.

A soma destes tempos faz com que entre a chegada de uma questão ao sistema e a devolução da resposta decorra, mais ou menos, uma semana. Nos casos urgentes este tempo pode ser reduzido a dois ou três dias, podendo chegar mesmo a meio dia.

— *Desenvolvimentos futuros*

Vários projetos estão em estudos ou em vias de realização, visando otimizar o trabalho do sistema. Podem ser citados:

1) Recorrer à literatura primária em substituição à secundária, reduzindo pela metade o tempo de entrada.

2) A perspectiva de aquisição de memórias de acesso direto maiores (IBM 2314), permitindo a utilização de arquivos inversos que diminuem consideravelmente o tempo de pesquisa em computador, atualmente efetuado segundo o princípio de arquivo direto (cada documento é seguido de seu conjunto de palavras-chave).

4.5. ANEXO 5 — O STAIRS

O STAIRS — STorage And Information Retrieval System — é um sistema que permite criar, atualizar e explorar bases de dados.

Divide-se em dois grandes grupos de funções: funções relativas à criação e manutenção das bases de dados e funções relativas à pesquisa e recuperação de informações.

Estas últimas são agrupadas sob o nome genérico de AQUARIUS — A QUery And Retrieval Interactive Utility System — que permite um diálogo entre o usuário e o sistema.

A recuperação de informações de uma base de dados STAIRS é efetuada submetendo ao sistema uma questão estruturada de acordo com critérios especificados pelo usuário.

A resposta a estas questões são unidades informativas, denominadas documentos, que satisfazem as condições impostas pela questão.

A obtenção de respostas satisfatórias dependerá basicamente de 4 fatores:

- 1) A estruturação adequada dos documentos.
- 2) O uso apropriado das funções do sistema.
- 3) O estabelecimento de uma estratégia de busca que reflita, o melhor possível, as necessidades do usuário.
- 4) O conhecimento da área coberta pela base de dados.

Um documento STAIRS se estrutura em cinco níveis. Do geral para o particular temos: documento, parágrafo, sentença, linha e palavra. Além destes, um documento STAIRS pode ter um ou mais campos formatados nos quais atuam certas funções, conforme será visto mais adiante.

A estrutura de um documento STAIRS pode ser exemplificada conforme segue.

DOCUMENTO 000001	
	ANO = 1984 → campo formatado
TÍTULO	Tecnologia do açúcar e química da sacarose.
AUTOR	Fulano de Tal.
ENTIDADE	Escola Superior de Doces. → parágrafo
DESCRIPTOR	Açúcar, Sacarose, Tecnologia, Obtenção e Refino, Desenvolvimento
OBJETIVOS	Desenvolvimento de tecnologia no campo da obtenção do açúcar e da sua refinação, realizando, para isso, pesquisas básicas dentro das operações unitárias que estão envolvidas nesse processo.
linha →	
sentença →	Desenvolvimento de tecnologia de obtenção de derivados da sacarose.

As características mais importantes do STAIRS são:

- 1) Divisão do documento em parágrafos, sentenças e campos formatados, cuja importância está relacionada com a existência de operadores que atuam sobre cada um destes subitens do documento.
- 2) Concatenação de bases de dados que, entre outras vantagens, permite manusear mais facilmente grandes coleções, subdividindo-as em bases físicas menores, concatenadas logicamente, permitindo pesquisá-las como se fosse uma única base de dados.

3) Consultas em linguagem natural e formato livre, possibilitando efetuar consultas na mesma linguagem do texto. Entretanto, se for desejado utilizar uma linguagem documentária, basta indexar o documento, colocar os descritores num parágrafo específico e limitar a consulta àquele parágrafo.

4) Utilização de operadores lógicos, contextuais, relacionais e especiais, cujos exemplos serão vistos mais adiante.

5) Exibição em destaque dos termos empregados na estratégia de busca. Esta característica faz com que as palavras usadas na questão sejam exibidas com luminosidade mais intensa que as outras.

6) Definição de sinônimos. Muito importante, principalmente ao trabalhar com linguagem natural. Esta característica permite "amarrar", entre si, vários termos sinônimos ou quase sinônimos, podendo utilizar qualquer um para recuperar ele e os demais.

7) Exibição e/ou impressão de determinadas partes do documento obtidas como resposta a uma determinada questão.

8) Respostas ordenadas por pertinência, obtidas utilizando uma função com 5 algoritmos diferentes para estabelecer esta pertinência.

9) Retrospecto da sessão. A qualquer momento podem ser verificadas as questões formuladas desde o início da sessão corrente.

10) Consultas "in batch". Isto quer dizer que lotes de consultas previamente preparados e testados podem ser submetidos ao sistema de forma não-conversacional. Esta facilidade tem grande aplicabilidade nas DSI, onde perfis de usuários, já testados, podem ser submetidos ao sistema de uma única vez e em horários em que o computador esteja menos concorrido.

Além destas existem muitas outras características importantes como, por exemplo:

Emissão de relatórios de contabilização.

Função de auxílio durante as sessões.

Possibilidade de armazenar consultas já efetuadas.

Obtenção de respostas ordenadas pelo conteúdo de um campo formatado.

Estabelecimento de níveis de privacidade.

Definição de palavras "vazias", para não poluir o dicionário da base de dados, etc.

Os recursos de que o STAIRS dispõe para formular questões são: operadores, qualificações e uso de parênteses e apóstrofes.

— Operadores

Os operadores se dividem em quatro grandes categorias:

operadores lógicos ou booleanos,

operadores contextuais,

operadores relacionais,

operadores especiais.

Operadores lógicos ou booleanos

- AND (a AND b) — Permite recuperar documentos contendo as palavras “a” e “b”, simultaneamente.
- OR (a OR b) — Permite recuperar documentos contendo a palavra “a” ou a palavra “b” ou ambas.
- XOR (a XOR b) — Permite recuperar documentos contendo a palavra “a” ou a palavra “b”, mas não ambas.
- NOT (a NOT b) — Permite recuperar documentos contendo a palavra “a”, mas não a palavra “b”.

Operadores contextuais

- ADJ (a ADJ b) — Recupera documentos contendo as palavras “a” e “b” adjacentes, nesta ordem.
- ADJn (a ADJn b) — Recupera documentos contendo as palavras “a” e “b” adjacentes, com até “n” palavras entre elas.
- SYN (a SYN b) — Considera “a” e “b” sinônimos na consulta corrente.
- NEAR (a NEAR b) — Recupera documentos contendo as palavras “a” e “b” adjacentes, em qualquer ordem.
- NEARn (a NEARn b) — Recupera documentos contendo as palavras “a” e “b” adjacentes, em qualquer ordem e com até “n” palavras entre elas.
- WITH (a WITH b) — Recupera documentos contendo as palavras “a” e “b”, desde que estejam na mesma sentença.
- NOTWITH (a NOTWITH b) — Recupera documentos que contenham as palavras “a” e “b”, desde que não estejam na mesma sentença.
- SAME (a SAME b) — Recupera documentos contendo as palavras “a” e “b”, desde que estejam no mesmo parágrafo.

• NOTSAME (a NOTSAME b) — Recupera documentos contendo as palavras “a” e “b”, desde que não estejam no mesmo parágrafo.

Operadores relacionais

Estes operadores se aplicam a campos formatados. Supondo um campo formatado denominado DATA, contendo o ano de publicação de uma obra, examinemos o uso destes operadores relacionais:

• EQ — (DATA EQ 1989) — Recupera documentos com data de publicação igual a 1989.

• NE — (DATA NE 1989) — Recupera documentos com data de publicação diferente de 1989.

• GT — (DATA GT 1989) — Recupera documentos com data de publicação posterior a 1989.

• NG — (DATA NG 1989) — Recupera documentos com data de publicação igual ou anterior a 1989.

• LT — (DATA LT 1989) — Recupera documentos com data de publicação anterior a 1989.

• NL — (DATA NL 1989) — Recupera documentos com data de publicação igual ou posterior a 1989.

• WL — (DATA WL 1984, 1989) — Recupera documentos com data de publicação entre 1984 e 1989, inclusive.

• OL — (DATA OL 1984, 1989) — Recupera documentos com data de publicação anterior a 1984 e posterior a 1989.

Operadores especiais

• MASK — Recupera documentos que contenham palavras iniciando por um radical dado. Exemplo: MASK COMPUT.

• ROOT — Funciona da mesma forma que o MASK, porém não produz nenhuma lista de documentos-resposta. É usado apenas para verificar no dicionário a presença ou não de determinada palavra e suas próximas.

• \$ — Semelhante ao MASK, pode ser usado dentro de uma questão. Exemplo: COMPUT\$

• * — Usado para ativar a função de sinonímia.

— *Qualificações*

- U/C — Recupera documentos onde a palavra dada esteja toda em maiúsculo (U/C indica Upper Case).
- L/C — Recupera documentos onde a palavra dada esteja toda em minúsculo (L/C indica Lower Case).
- M/C — Recupera documentos contendo a palavra dada com caracteres maiúsculos e minúsculos, pelo menos um de cada (M/C indica Mixed Case).
- F/C — Recupera documentos onde a palavra dada possua o caráter inicial em maiúsculo (F/C indica First Case).
- N/P — Recupera documentos desde que a palavra dada esteja presente no parágrafo especificado (N/P indica Name ou Number Paragraph). Pode ser especificado também na forma negativa, ou seja, a palavra não deve estar presente no parágrafo especificado.

— *Parênteses e apóstrofes*

Os parênteses são usados para cercar as partes da questão a serem avaliadas independentemente. Em outras palavras: são usados para alterar a hierarquia dos operadores.

Os apóstrofes são usados para indicar que o que está entre eles é uma palavra a ser recuperada. Por exemplo se quisermos recuperar documentos contendo a palavra “adj”, devemos colocar o “adj” entre apóstrofes — ‘adj’, para que não seja confundido com o operador ADJ.

4.6. ANEXO 6 — INFORMÁTICA E CULTURA: UM EXEMPLO DE PIONEIRISMO

Entre os programas desenvolvidos pelo Instituto Cultural Itaú, constituído em fevereiro de 1987, dois deles merecem ser citados aqui, pelos vínculos que têm com a biblioteconomia e a documentação, bem como com a informática.

Trata-se, de fato, de dois programas de amplo alcance nos quais, utilizando ferramentas da informática mais sofisticada, alimentam-se e exploram-se bancos de dados contendo informações sobre artes em geral.

Um destes bancos se compõe de bases de dados contendo informações factuais (biografias, produção artística e dados histórico-críticos) sobre pintores brasileiros ou relativos ao Brasil, dos séculos XIX e XX, devendo evoluir até abranger todas as artes plásticas e posteriormente as demais artes.

Associado a este banco de dados existe outro, contendo imagens digitalizadas da produção artística mais representativa destes pintores.

O banco de dados do segundo programa é composto de informações bibliográficas em artes, sem limites temporais ou geográficos, devendo evoluir abrangendo outros tipos de suportes da informação, tais como filmes, discos, vídeos, diapositivos e assim por diante.

Ambos os bancos de dados são administrados pelo software documentário STAIRS — SStorage And Information Retrieval System — apresentado no anexo 5.

O interesse destes programas, pioneiros em nosso meio, para os profissionais da informação (que é o que bibliotecários e documentalistas são, ou deveriam ser) é por demais óbvio para ser justificado. Entretanto, apenas como comentário, vale a pena chamar a atenção sobre o enorme campo de trabalho que se descortina para os que lidam com informações de todos os gêneros, registradas sobre os mais diversos e sofisticados veículos.

Este imenso campo de atuação, do qual estes programas pioneiros são exemplo, implica, evidentemente, um preparo dos profissionais que deve ir além dos currículos acadêmicos, dos preconceitos e dos convencionalismos confortáveis.

Já se foi o tempo em que o bibliotecário era preparado como se fosse atuar na biblioteca de Alexandria, lidando com palimpsestos e incunábulo e tendo como usuários monges e sábios que, entre outras coisas, dispensariam perfeitamente seus serviços.

Estamos na era da informática e para participar da mesma não basta apenas sentar diante de um microcomputador qualquer e fazer dois ou três cursinhos de fim de tarde.

Para lidar com a informática é preciso, mais do que nunca, bom senso e espírito crítico, além dos conhecimentos técnicos.

A informática é uma arma de dois gumes, pois os erros também são processados com velocidade eletrônica e assim sendo eles fogem ao nosso controle num piscar de olhos.

Subestimar o computador é um erro. Superestimá-lo pode ser fatal.

Encerrado este comentário voltemos ao nosso exemplo.

O banco de dados com informações factuais compreende três bases de dados distintas, fisicamente separadas, mas unidas logicamente por uma estrutura comum.

Isto significa que cada base de dados pode ser pesquisada separadamente, mas, quando e se desejado, as três podem ser pesquisadas como se fossem uma única.

A estrutura possui itens, ou parágrafos, para usar a terminologia do STAIRS, dentro dos quais a informação global se distribui.

Assim sendo, temos parágrafos como:

Nome do artista

Data de nascimento

Data da morte

Filiação

Formação

Viagens

Alunos

Escolas/movimentos

Gêneros

Atividades artísticas

Outras atividades

Homenagens

Dados histórico/críticos

Dados da obra

Trabalhos prévios

Etc.

Estes parágrafos podem ser exibidos na sua totalidade ou em grupos, segundo o interesse específico do usuário.

Como já foi dito, estas bases de dados são gerenciadas pelo STAIRS. Entretanto, considerando a extrema heterogeneidade dos usuários, foi adotado um software, desenvolvido por uma equipe da área de sistemas e métodos do Banco Itaú, que funciona como interface entre as questões formuladas e as funções de recuperação do STAIRS, bem como entre os resultados obtidos pelo STAIRS e sua apresentação na tela e/ou no papel.

Este banco de dados possui um vínculo com o banco de dados bibliográfico, que descreveremos a seguir, e com o banco de dados de imagens digitalizadas de obras de vários artistas.

Assim sendo, é possível verificar, por um lado, as referências bibliográficas existentes sobre determinado artista ou assunto e, por outro lado, as imagens existentes de um determinado pintor, de uma determinada escola, de um determinado período.

Estas imagens, capturadas e transformadas em bits para serem guardadas em memórias magnéticas, podem ser exibidas nas telas de monitores e/ou impressas com uma fidelidade bastante satisfatória.

O banco de dados bibliográficos é constituído, no momento, por uma base de dados onde se procurará centralizar todo tipo de documento impresso que verse sobre arte, oriundo dos mais diversos acervos de bibliotecas e museus.

Esta diversidade de origens implica uma série de controles de duplicação e terminológicos, sem os quais a coleção se tornaria impraticável em pouco tempo.

Ao longo do tempo mais acervos serão incorporados, vindos de instituições de outros estados e até mesmo de outros países, assim como mais bases de dados, contendo informações veiculadas por outros suportes que não o documento impresso, serão criadas e adicionadas a esta existente e em vias de expansão.

Nesta base de dados são registradas informações reportadas por livros, periódicos, programas de óperas, catálogos de exposições, teses, roteiros de cinema, partituras, e assim sucessivamente. Daí a necessidade de uma estrutura que pudesse atender às características particulares de cada material, para evitar ter de criar tantas bases de dados quantos fossem os materiais impressos existentes, pois isto acarretaria a necessidade de formular a mesma questão em cada base, tornando o processo moroso e cansativo.

Assim sendo, nesta estrutura bastante grande, temos itens díspares como:

Autor

Título

Compositor

Artista

Diretor artístico

Curador
Orquestra
Regente da orquestra
Coro
Regente do coro
Tradutor
Prefaciador
Caricaturista
Ilustrador
Fotógrafo
Imprensa
Idioma
Coreógrafo
Bailarinos
Intérpretes
Música
Iluminador
Atos
Personagens
Arranjador
Autor do libreto
Notas
Descritores
Resumo
e muitos outros mais.

Da forma sucinta e horizontal como foi colocado neste anexo, pode parecer que estes dois programas são resultado de um trabalho simples e rápido. Nada mais errado.

Primeiro, há que se considerar o enorme e delicado trabalho de captura das informações, envolvendo especialistas na área, capazes de localizar, selecionar, agrupar, resumir e preparar estas informações para os processamentos posteriores.

Num segundo momento, entram no processo especialistas em informática, aos quais compete tornar estas informações disponíveis aos usuários, utilizando os mais modernos recursos da tecnologia computacional e afins.

Apenas para dar uma idéia das dimensões do trabalho, basta lembrar que o levantamento da biografia de um artista pode levar

semanas de penosas pesquisas, não esquecendo que as informações factuais são dinâmicas e requerem atualizações constantes, por força daquilo que elas refletem.

Da mesma forma, o controle e a uniformização da terminologia que vai gerar os descritores é um trabalho praticamente sem fim.

Não podemos deixar de salientar que a preparação conceitual da estrutura de uma base de dados complexa pode levar meses de trabalho, principalmente em casos onde uma mesma estrutura deve hospedar informações oriundas de origens as mais diversas. Isto sem contar que estas estruturas devem ser suficientemente flexíveis para poderem sofrer eventuais adaptações ditadas pelo uso ou feedback.

Assim sendo, a simplicidade desse trabalho pioneiro fica apenas por conta das aparências.

A finalidade deste anexo nos impede de detalhar convenientemente estes programas. Em todo caso eles estão aí para serem conhecidos e, quem sabe, em outra oportunidade possamos apresentá-los e discuti-los da forma que eles merecem.

4.7. ANEXO 7 — GLOSSÁRIO DE TERMOS, SIGLAS E ACRÔNIMOS

Neste glossário estão reunidos termos “informáticos” constantes desta publicação, incluindo siglas e acrônimos acompanhados de sua forma extensa.

Considerando nossa preocupação, mais voltada aos leitores pouco ou nada iniciados no assunto do que ao rigor semântico, pedimos a compreensão dos leitores mais exigentes para alguns conceitos pouco ortodoxos.

Analógico: Que respeita a formação do modelo real.// Diz-se de aparelhos que representam quantidades por meio de grandezas físicas.

AQUARIUS: A QUery And Retrieval Interactive Utility System (Sistema Utilitário e Interativo de Perguntas e Recuperação).

Arquivo: Coleção homogênea de dados, passível de ser memorizada e processada por um computador.

ASCII: American Standard Code for Information Interchange (Có-

digo Americano de Normalização para Intercâmbio de Informações).

Automação: Funcionamento de uma máquina ou conjunto de máquinas que, sob controle de um programa, torna possível a realização de tarefas, sem intervenção humana no decorrer do processo.

Batch: Lote. Batch Processing — Processamento por lotes de dados.

BCD: Binary-Coded Decimal (Binário Codificado em Decimal).

Binário: Conjunto que comporta dois elementos, dois dispositivos ou dois sistemas. Por extensão: sistema de numeração de base 2 que permite realizar operações com apenas dois algarismos.

BIT: BInary digiT (Dígito binário).

Booleano: Relativo à lógica matemática binária de Georges Boole.

BPS: Bits Por Segundo (velocidade de transmissão).

Byte: Conjunto de 8 bits utilizado para representar um símbolo ou um caráter.

Campo: Parte de um registro contendo uma informação definida.

CD-ROM: Compact Disk-Read Only Memory (Disco Compacto, Usado Apenas para Leitura).

CDU: Classificação Decimal Universal.

CID: Centre of Information and Documentation (Centro de Informação e Documentação).

COM: Computer Output Microform (Saída de Computador por Microforma).

Compilador: Programa tradutor de programas escritos em linguagem evoluída para programas em linguagem executável.

Consistência: Verificação efetuada por um programa no sentido de assegurar que os dados a serem processados estejam em conformidade com as especificações.

CPD: Centro de Processamento de Dados.

CPU: Central Processing Unity (Unidade Central de Processamento).

Dado: Informação passível de entrar num sistema de processamento e ser explorada para um cálculo aritmético ou lógico.

Decimal: Sistema de numeração de base 10.

Digital: Que representa dados por meio de dígitos.

- Dígito:** Posição de contagem dentro de um número. // Algarismo.
- DSI:** Disseminação Seletiva da Informação.
- Disk Pack:** Pilha de discos magnéticos, montados num único eixo, de grande capacidade de armazenamento, utilizada em computadores de grande porte.
- Disquete:** Disco magnético, flexível, removível, de capacidade de armazenamento relativamente pequena.
- EBCDIC:** Extended Binary-Coded Decimal Interchange (Binário Codificado em Decimal, para Intercambio).
- ENDS:** European Nuclear Documentation System (Sistema Europeu de Documentação Nuclear).
- ENIAC:** Electronic Numeric Integrator And Calculator (Computador e Integrador Numérico Eletrônico).
- EPROM:** Erasable Programmable Read Only Memory (Memória Programável e Apagável, Somente para Leitura).
- Fluxograma:** Representação gráfica de seqüência de procedimentos.
- Hardware:** Todo dispositivo físico que constitui um sistema de computador. // Maquinaria.
- Hexadecimal:** Sistema de numeração de base 16.
- ILS:** Integrated Library System (Sistema Integrado de Biblioteca).
- Informação:** Conceito abrangendo toda e qualquer formulação abstrata que permita refletir e descrever um elemento — situação, propriedade, acontecimento — de um sistema estruturado qualquer.
- Informática:** Na acepção usada nesta publicação — estudo dos meios de processamento da informação. (Esta acepção, da linha franco-russa, difere da acepção americana pela qual informática é o mesmo que computação.)
- INIS:** International Nuclear Information System (Sistema Internacional de Informações Nucleares).
- Instrução:** Ação de dar ordem a uma máquina ou parte dela.
- IPEN:** Instituto de Pesquisas Energéticas e Nucleares.
- ISO:** International Standard Organization (Organização Internacional de Normalização).

- Dígito:** Posição de contagem dentro de um número. // Algarismo.
- DSI:** Disseminação Seletiva da Informação.
- Disk Pack:** Pilha de discos magnéticos, montados num único eixo, de grande capacidade de armazenamento, utilizada em computadores de grande porte.
- Disquete:** Disco magnético, flexível, removível, de capacidade de armazenamento relativamente pequena.
- EBCDIC:** Extended Binary-Coded Decimal Interchange (Binário Codificado em Decimal, para Intercambio).
- ENDS:** European Nuclear Documentation System (Sistema Europeu de Documentação Nuclear).
- ENIAC:** Eletronic Numeric Integrator And Calculator (Computador e Integrador Numérico Eletrônico).
- EPROM:** Erasable Programable Read Only Memory (Memória Programável e Apagável, Somente para Leitura).
- Fluxograma:** Representação gráfica de seqüência de procedimentos.
- Hardware:** Todo dispositivo físico que constitui um sistema de computador. // Maquinaria.
- Hexadecimal:** Sistema de numeração de base 16.
- ILS:** Integrated Library System (Sistema Integrado de Biblioteca).
- Informação:** Conceito abrangendo toda e qualquer formulação abstrata que permita refletir e descrever um elemento — situação, propriedade, acontecimento — de um sistema estruturado qualquer.
- Informática:** Na acepção usada nesta publicação — estudo dos meios de processamento da informação. (Esta acepção, da linha franco-russa, difere da acepção americana pela qual informática é o mesmo que computação.)
- INIS:** International Nuclear Information System (Sistema Internacional de Informações Nucleares).
- Instrução:** Ação de dar ordem a uma máquina ou parte dela.
- IPEN:** Instituto de Pesquisas Energéticas e Nucleares.
- ISO:** International Standard Organization (Organização Internacional de Normalização).

Software: Todo trabalho intelectual relacionado ao funcionamento de computadores. // Programas.

STAIRS: STorage And Information Retrieval System (Sistema de Armazenamento e Recuperação de Informações).

Terminal: Máquina orientada para a comunicação de dados, geralmente situada na proximidade da origem dos mesmos.

UAL: Unidade Aritmética e Lógica.

UC: Unidade de Controle.

UCP: Unidade Central de Processamento.

UCT: Unidade de Controle de Transmissão.

UMC: Unidade de Memória Central (ou Principal).

USASI: United States of America Standard Institute (Instituto de Normalização dos Estados Unidos da América do Norte).

Winchester: Disco magnético, rígido, fixo, de capacidade relativamente grande, utilizado em microcomputadores.

4.8. ANEXO 8 — MINILÉXICO INGLÊS-PORTUGUÊS

Access	Acesso
Accounting	Contabilidade
Accumulator	Acumulador
Accuracy	Exatidão
Actual address	Endereço absoluto (de posição de memória)
AND	E (operador lógico)
Asynchronous	Assíncrono
Attach (to)	Associar a // Conectar
Available	Disponível
Backspace	Espaço atrás
Batch	Lote
Binary	Binário
Branch	Desvio
Call (to)	Chamar
Card	Cartão

Software: Todo trabalho intelectual relacionado ao funcionamento de computadores. // Programas.

STAIRS: STorage And Information Retrieval System (Sistema de Armazenamento e Recuperação de Informações).

Terminal: Máquina orientada para a comunicação de dados, geralmente situada na proximidade da origem dos mesmos.

UAL: Unidade Aritmética e Lógica.

UC: Unidade de Controle.

UCP: Unidade Central de Processamento.

UCT: Unidade de Controle de Transmissão.

UMC: Unidade de Memória Central (ou Principal).

USASI: United States of America Standard Institute (Instituto de Normalização dos Estados Unidos da América do Norte).

Winchester: Disco magnético, rígido, fixo, de capacidade relativamente grande, utilizado em microcomputadores.

4.8. ANEXO 8 — MINILÉXICO INGLÊS-PORTUGUÊS

Access	Acesso
Accounting	Contabilidade
Accumulator	Acumulador
Accuracy	Exatidão
Actual address	Endereço absoluto (de posição de memória)
AND	E (operador lógico)
Asynchronous	Assíncrono
Attach (to)	Associar a // Conectar
Available	Disponível
Backspace	Espaço atrás
Batch	Lote
Binary	Binário
Branch	Desvio
Call (to)	Chamar
Card	Cartão

Gap	Intervalo
Half	Metade//Semi
Hardware	Maquinaria//Parte física de um computador
Header	Cabeçalho//Inicial//Na frente
High-speed	Rápido// Alta velocidade
In-line	Imediato
Index	Índice
Input	Entrada (de dados)
Inquiry	Interrogação
Instruction	Instrução
Inter-block gap	Intervalo entre registros (blocos)
Jump	Salto
Key	Chave//Tecla//Interruptor
Key board	Teclado
Label	Rótulo//Etiqueta
Latency	Tempo de espera
Leading zeroes	Zeros não significativos//Zeros à esquerda
Library program	Biblioteca de programas
Linkage	Concatenação
List (to)	Listar
Load (to)	Carregar
Location storage	Ponto de memória
Loop	Laço//Circuito de instruções
Main	Principal
Maintenance	Manutenção
Master	Matriz//Mestre//Principal
Medium	Meio//Suporte
Merge	Intercalação
Mistake	Erro
Network	Rede de circuitos
NOT	NÃO (operador lógico)
Notation	Notação//Numeração

Off-line	Fora da linha//Autônomo//Independente
On-line	Em linha//Conectado
OR	OU (operador lógico)
Output	Saída
Overflow	Ultrapassagem da capacidade (de gravação)
Pack (to)	Compactar//Agrupar
Parity bit	Bit de paridade//Bit de teste
Plot (to)	Traçar (curvas)
Printer	Impressora
Priority	Prioridade (de execução, por exemplo)
Procedure	Procedimento
Processing	Processamento
Punch (to)	Perfurar (cartões, fitas de papel)
Random	Aleatório//Randômico//Direto
Reader	Leitora (dispositivo de entrada de dados)
Record	Registro
Relocable	Transferível
Report	Relatório
Reset	Restaurar//Reestabelecer//Zerar
Restart	Recomeçar (dar novo início)
Restore	O mesmo que Reset
Routine	Rotina//Programa
Run	Rodada//Passagem (de um programa pela máquina)
Screen	Tela//Vídeo (de terminal)
Search	Pesquisa//Procura//Busca
Seek	Posicionamento (de cabeçote de leitura/gravação)
Select (to)	Selecionar
Sequential	Seqüencial
Size	Tamanho//Dimensão
Software	Produção intelectual associada a computadores
Sort	Ordenação//Classificação//Alfabetação
Space	Espaço
Speed	Velocidade
Step	Passo (de programa)

Store (to)	Memorizar//Armazenar
String	Cadeia//Fila
Switch	Chave//Comutador//Interruptor
Tag	Etiqueta//Rótulo
Tape	Fita (perfurada ou magnética)
Task	Tarefa
Test	Teste//Ensaio//Prova
Track	Trilha//Pista (de gravação)
Underflow	Valor inferior ao próprio limite
Unit	Unidade//Dispositivo//Equipamento//Periférico
Unpack (to)	Descompactar
Update (to)	Atualizar
User	Usuário
Utility (program)	Programa utilitário
Validity check	Controle de validade//Teste de validade
Waiting time	Tempo de espera (de resposta do computador)
Word	Palavra (de máquina)
Write (to)	Escrever//Gravar (em memória)
Writing head	Cabeçote de gravação
Zone	Zona

4.9. ANEXO 9 — UMA INCURSÃO NO AMANHÃ (À GUIA DE CONCLUSÃO)

Levando-se em conta a situação atual e os estudos e pesquisas em curso, podemos antever os caminhos que serão trilhados pela informática nos próximos anos.

— A colocação de textos completos nas memórias dos computadores poderá ser feita por equipamentos especiais que transformarão os textos em seqüências de bits aceitáveis por estas máquinas, evitando a transcrição com todos os problemas a ela inerentes. Estes

textos, uma vez memorizados, poderão ser processados por programas especiais que os coloquem nos formatos desejados.

— Haverá um aumento muito grande na capacidade de armazenamento das memórias auxiliares, quer seja em meios magnéticos, óticos ou sonoros.

— Como decorrência dos dois itens acima, teremos enormes bancos de dados contendo textos completos, complementados por imagens, gráficos e sons.

— Estes bancos de dados poderão ser acessados por softwares realmente “amigáveis ao usuário”, no sentido de não exigirem nenhum conhecimento “informático” prévio, aceitando as formulações na linguagem do próprio usuário e efetuando as pesquisas nos textos integrais com o auxílio de sistemas peritos (experts) interagindo nas formulações.

— O diálogo entre o homem e a máquina será profundamente modificado, não apenas pela utilização de softwares “amigáveis”, mas também pela utilização de terminais de comunicação diferenciados, em substituição, ou como alternativa, ao velho vídeo de tubo catódico.

— O grau de integração dos circuitos integrados poderá ser multiplicado por 20, ou mesmo 30, até o final do século, resultando num aumento de memória que poderá chegar perto de 100 milhões de bits.

— A potência dos computadores será multiplicada. Enquanto os processadores individuais parecem limitados a cerca de 100 M Flops (1 M Flop = 1 milhão de operações flutuantes por segundo), estruturas paralelas deverão aparecer nos próximos anos, levando os M Flops a beirar os 10.000.

— As possibilidades de comunicação entre computadores, realizando o intercâmbio de informações, crescerão graças à vulgarização de redes telefônicas, satélites, e outros meios mais, que tornarão bastante simples a conexão de várias fontes informáticas sobre uma única rede. Neste processo, o desenvolvimento das fibras óticas terá papel preponderante.

— Se a tudo isto acrescentarmos jornais eletrônicos, disseminação maciça do videotexto (ou similares), terminais baratos, eficientes e portáteis, auto-educação via computadores, e outras “maravilhas” mais, cabe uma pergunta: qual será o papel e o perfil do profissional da informação do século XXI???

SOBRE O AUTOR

Nascido em Pisa, Itália, cheguei ao Brasil aos treze anos de idade. Anos mais tarde naturalizei-me brasileiro para poder participar da vida política do país.

Formado em Biblioteconomia e Documentação pela ECA/USP, enveredei pelos caminhos da informática, sem, contudo, distanciar-me jamais da minha área de formação.

Entrei no curso de Filosofia da USP, interrompendo-o no final do segundo ano por absoluta falta de tempo (trabalhava de manhã e cursava a ECA à tarde).

Imediatamente após formado trabalhei no Centro de Processamento de Dados do então Instituto de Energia Atômica — IEA — na parte de automação da biblioteca daquela instituição. Pelo IEA estagiei, durante 6 meses, no European Nuclear Documentation System da Comunidade Européia, sediado em Luxemburgo, para, na volta, organizar e administrar bases de dados contendo informações nucleares.

Saí do IEA, sendo convidado para trabalhar no Departamento de Ciência e Tecnologia da Secretaria da Indústria, Comércio, Ciência e Tecnologia do Estado de São Paulo, para gerenciar bases de dados sobre instituições de pesquisa, pesquisadores e pesquisas em andamento no Estado de São Paulo, indo assessorar os estados da Bahia, Ceará e Paraná, em trabalhos semelhantes.

Lecionei informática na ECA e na FATEA durante vários anos.

Posteriormente atuei como analista de sistemas do Banco Itaú e atualmente coordeno o Centro de Referência Bibliográfica do Instituto Cultural Itaú.

Publiquei algumas monografias técnicas e escrevi alguns artigos em revistas especializadas.

Impresso na
Gráfica A Tribuna de Santos Ltda.
Rua João Pessoa, 349
Telefone 32.8692
CEP 11.013 - Santos -

A *Coleção Palavra-chave* tem por objetivo oferecer aos profissionais bibliotecários, aos estudantes de Biblioteconomia e Documentação e a outros interessados, textos básicos e acessíveis sobre temas relevantes e atuais relacionados com o campo da Biblioteconomia, Documentação e Ciência da Informação.

A quase inexistência de textos nessas áreas exigia das editoras e entidades um trabalho conjunto visando preencher essa lacuna. A *Coleção Palavra-chave* surge no bojo desse esforço e orgulha-se em ser a primeira com a preocupação de editar e veicular textos que contribuem para a atualização e aperfeiçoamento dos profissionais que atuam no âmbito da Biblioteconomia, Documentação, Ciência da Informação e áreas afins.

A escolha do texto *A Informática na Biblioteconomia e na Documentação* para fazer parte da *Coleção palavra-chave*, deve-se à atualidade do assunto, ao interesse e preocupação que desperta nos profissionais que atuam na área e, principalmente, pela importância que estão assumindo as novas tecnologias da informação no ambiente bibliotecário. O tema é atual e adequado às necessidades de relevante parcela de profissionais bibliotecários, pois contribui para elucidar aspectos básicos de um instrumental que, cada vez mais, vem exigindo maior espaço e impondo-se como imprescindível.

editora polis
associação paulista de bibliotecários