

1

ANAIS DO I CONGRESSO
BRASILEIRO DE ARQUIVOLOGIA
01



ASSOCIAÇÃO DOS ARQUIVISTAS BRASILEIROS

**ANAIS DO I CONGRESSO
BRASILEIRO DE ARQUIVOLOGIA**

Rio de Janeiro, de 15 a 20 de outubro de 1972

O ARQUIVO E O COMPUTADOR

Antônio Garcia de Miranda Netto

Considera a origem da palavra arquivo, emite os conceitos de arquivo, computador e informática e sugere o estabelecimento de um glossário comum a todos os campos da informação, como meio de evitar uma terminologia ambígua. Descreve o computador a título de apresentação. Aborda considerações de natureza técnica, quanto ao funcionamento do computador, e indica terminologia informática, comparando-a à utilizada na Arquivologia. Analisa memória e arquivo no computador e estabelece a comparação do sentido da palavra arquivo no computador e na conotação clássica em que é utilizada. Formula considerações quanto a tratamento de dados em computadores e arquivos. Por fim, enfatiza a necessidade inadiável dos especialistas nos campos da Arquivologia, Biblioteconomia, Bibliobiologia, Museologia, Documentação, Análise de Sistemas, Informática, Administração e Direito, reunidos trabalharem para o estabelecimento de uma tecnologia avançada.

THE ARCHIVE AND THE COMPUTER

The author discusses the origins of the word "archive" and the concept of archive, computer and informational system, and suggests the elaboration of a glossary for use in all fields of information in order to avoid ambiguous terminology. He describes the computer and informational technology, comparing it to that used in archival science. He analyzes the memory and archive of the computer, comparing the meaning of the word 'archive' in the computer with its traditional meaning. He also discusses the treatment of data in computers and archives. He emphasizes the urgent need for specialists to try to understand the archive, since all its theoretical and practical aspects must be linked to Library Science, bibliobiology, museumology, documentation, systems analysis, administration and law, all working together to establish advanced technology.

L'ARCHIVE ET LE CALCULATEUR

Il considère l'origine du mot: archive; l'exprime les idées d'archive: calculateur et informateurs et il suggère l'établissement d'un glossaire commun à toutes es matières de renseignement, comme moyen d'éviter une terminologie équivoque. Il décrit le calculateur comme une présentation. Il aborde des considérations de nature technique, à l'égard du fonctionnement du calculateur, et montre une terminologie informatrice en la comparant à celle dont on se sert dans l'archivologie. Il analyse la mémoire et l'archive dans le calculateur et établit une comparaison entre

le sens du mot, archive dans le calculateur et dans l'expression classique dant an s'est servi. Il formule des considérations sur l'emploi de données en calculateurs et archives. Enfin il fait ressortir la nécessité qui ne peut être ajournée des spécialistes de s'unir en travaillant dans: l'Archivologie, Bibliothéconomie, Bibliologie, Muséologie, Documentation, Analyse de Systèmes Informatrice, Administration et Droit, pour l'établissement d'une technologie avancée.

A Professora Marilena Leite Paes ARIADNE cujas lições foram o fio que me guiou pelos labirintos da Arquivologia

SUMÁRIO

Introdução — 1. Apresentação do Computador — 2. Anatomia e Fisiologia do Computador — 3. Terminologia Informática e Arquivologia — 4. Memória e Arquivo no Computador — 5. Arquivo de Computador e Arquivo Convencional — 6. Lógica do Processamento — 7. Tratamento de Dados em Computadores e Arquivos — 8. Necessidade de um Trabalho Conjunto — Bibliografia.

Introdução

Há uma relação estreita — muito mais do que se possa imaginar à primeira vista — entre o analista de sistemas-computador, o especialista em arquivos, o documentalista e o usuário esclarecido (instituição ou pessoa).

A chamada Revolução Informática foi o resultado de uma transformação radical no próprio conceito de ciência e suas aplicações. A passagem de um modelo estático do existente, como um conjunto de categorias estanques, para um modelo dinâmico de sistema geral, composto de subsistemas intimamente ligados, levou à substituição da diáde Energia/Matéria, dominante a partir do Século XVIII e até meados do nosso, por outro binômio: Informação/Ação. (¹) A ação é a base de qualquer apropriação do real, a partir de um projeto de pensamento. Mas ninguém começará a pensar sem dados e fatos, e o projeto nada mais é que uma ordenação finalística de dados e fatos. A INFORMAÇÃO passa, pois, ao primeiro plano como o principal parâmetro condicionador de qualquer atividade humana. E não será simples cumprimento de um generalista, feito aos especialistas nacionais, reunidos neste PRIMEIRO CONGRESSO BRASILEIRO DE ARQUIVOLOGIA, a afirmação de que o ARQUIVO, no sentido mais lato, é base e fundamento das três atividades humanas: Contemplar, Experimentar e Agir. Os três aspectos consequentes da criação, o teórico, o empírico e o prático estão sempre intimamente ligados aos documentos (dados e fatos) (²) e suas correlações.

Dados e fatos — registrados na memória de cada um ou em suporte material — são os elementos que irão formar a informação elaborada,

base da ação. Não seria nenhuma impropriedade chamar de ARQUIVO a qualquer coleção de dados ou fatos que vise a enriquecer um repertório de conhecimentos, seja para construir um projeto de ação e transformar o existente, seja para maior enriquecimento interior. Assim a Biblioteca, a Pinacoteca, uma coleção de Manuais de Serviço, de Leis e Regulamentos ou a Programoteca de um sistema-computador podem perfeitamente ser denominados ARQUIVOS, sem que em nada forcemos a palavra.

O primeiro sentido de “arquivo”, em português, é o registrado por Frei Domingos Vieira no início do século passado: “lugar em que se guardão títulos, documentos escripturas de uma família, corporação e principalmente nação”; aos poucos foi tomando novas conotações, e tem vários sentidos diferentes.

A palavra “arquivo” deriva do latim *archivum*, voz de baixa latinidade que só apareceu no Império Romano no século II da nossa era, na forma *archium* registrada em Suetônio (historiador que viveu de 69 a 125), mudada no século III para a forma final *archivum* pelo jurisconsulto Ulpiano (nascido em Tiro, na Fenícia, e morto em Roma em 228), assimilação tardia do grego *arkheion*, palácio do tribunal, com o mesmo sentido. O lugar onde se guardavam os documentos se dizia, em grego, *grammatophylakeion*, e o arquivista era o *grammatophylaks* (guardião das causas escritas). No latim clássico o arquivista era *chartarius*, *chartularius* (forma que predominou na Idade Média) ou *chartophylax*, decalque do grego. (³)

Este brevíssimo excuso visa apenas a mostrar que não devemos dar valor absoluto às palavras, que vão e vêm com o passar do tempo, desaparecendo ou tomando novas acepções. Mas não poderemos pensar sem palavras, e a necessidade de uma terminologia não ambígua em cada disciplina, pelo menos durante certo período de tempo, é inegável. Não tenhamos a ilusão de que é permanente qualquer terminologia, embora os termos técnicos sejam menos sujeitos a transformações que os da linguagem quotidiana. A terminologia não ambígua é indispensável instrumento de comunicação e ação. Seria pois da maior utilidade o estabelecimento de um glossário comum a todos os campos da Informação e Documentação (incluindo procedimentos e técnicas de microfilme, em particular, e reprografia em geral). Esta comunicação visa a mostrar a necessidade de uma estreita colaboração entre especialistas dos três campos: Informática, Documentação e Arquivo.

1. Apresentação do Computador

O especialista em arquivos terá uma visão mais completa de seu campo se conhecer o papel que representam os modernos dispositivos eletrônicos no tratamento da informação. O não iniciado terá uma sensação de surpresa ao ver, pela primeira vez, um computador. O aspecto da máquina não lhe dá, imediatamente, a idéia de sua capacidade. Como o avião supersônico ou o foguete espacial, o computador é um “engenho

com finalidade". Mas o Concorde, ao vencer a barreira do som, ou a Apolo 15 subindo no fragor dos jatos de fogo, dão logo a impressão perfeita de sua força e finalidade. O tear automático e o torno programado mostram logo o que estão fazendo, de modo tão perfeito, sem intervenção direta do homem. Já o computador não nos impressiona senão em um segundo tempo, depois que nos são reveladas as múltiplas tarefas que pode desempenhar. O próprio produto acabado do computador (informação), embora esteja em suporte material, não tem uso material: não se *usa* a informação como se usa um par de sapatos ou um aparelho de barbear elétrico. Também o trabalho do computador (processamento) se faz sem ruído e sem movimento no interior do conjunto de pequenos núcleos magnetizáveis, que constituem a memória central da Unidade Central de Processamento.

O computador é uma máquina-ferramenta. Como toda máquina-ferramenta, recebe matéria-prima e a transforma em produto final, através de dispositivos e mediante um programa e um operador. Esta transformação se chama *processo*. Mas não se pense que o ciclo "matéria-prima-processo-produto acabado" exija sempre mecanismos complicados. O simples gesto de ajeitar um laço no vestido ou aparar um lápis apresenta as mesmas fases de entrada de matéria-prima, processo e saída do produto acabado.

O computador é um instrumento de múltiplos talentos – versátil como hoje se diz – capaz de cumprir as mais variadas tarefas e tramar as combinações mais impressionantes. Sua matéria-prima são os fatos e dados que vão ser transformados em "informação". Por isso se chama INFORMÁTICA (outro termo em moda, de que muito se abusa) ao "tratamento" (ou processamento) automático e racional considerado como base do conhecimento e da comunicação, em geral com o uso de mecanismos complexos denominados calculadoras eletrônicas, computadores ou, ainda, sistemas informáticos. Seria mais correto dizer processamento de dados e não de informações, pois a informação é o produto final.

Fato ou dado é algo que pode ser descrito, situado ou medido, em linguagem não equívoca ou em escala de medida de espaço, tempo ou intensidade (como o sistema CGS, por exemplo). Em terminologia de computador, *dado* significa "um grupo de operandos ou fatores constituídos de algarismos, caracteres alfabéticos ou símbolos, que denotam qualquer condição, estado ou valor dos fatos tratados em um programa, mas não o próprio programa".⁽⁵⁾ O uso do computador acentuou a diferença entre "dados" e "informação". Informação é "reunião, combinação, análise ou resumo dos dados, em forma comprehensível e utilizável".⁽⁶⁾

Poderá alguém dizer que a informação já estava nos fatos e que o próprio fato já é uma informação. Mas há um certo matiz de diferença entre as duas situações. Atoine Coysevox, o grande escultor da época de Luís XIV, talhava um bloco de mármore quando dele se aproxima um dos cortezãos do Rei-Sol que lhe pergunta se estava fazendo um cavalo. "Não – responde o artista – estou *descobrindo* um cavalo". O cavalo estava no

mármore, como a informação nos fatos, mas foi preciso que alguém o mostrasse. Fatos são tijolos — como dizia Lord Kelvin — e nada significam em si mesmos. Só começam a ter sentido quando se tornam elementos de uma construção, que é a hipótese ou a teoria.

Essa admirável máquina-ferramenta que transforma os dados em informação não é tão complicada como poderia parecer à primeira vista. Ainda há quem a chame de cérebro eletrônico” — como o faziam os otimistas dos primeiros tempos, que superestimavam a capacidade de combinação do computador e viam um substituto da inteligência humana em um dispositivo que não é mais que máquina. A característica essencial do computador é depender de um programa: não pode trabalhar sem instruções precisas, minuciosas e logicamente completas. Sua “inteligência” pode comparar-se à do jardineiro luso a quem dizia a patroa: “Manuel, chegará daqui a pouco uma senhora para visitar-me. Avise-a que tive de ir, com urgência, à cidade mas não demorarei mais de um quarto de hora. Faça-a entrar e diga-lhe que a criada a servirá no que desejar, pedindo-lhe que me espere.” E quando o automóvel já estava quase no portão, eis que vem o Manuel correndo e esbaforido: “Minha senhora, minha senhora! E se ela não vier, o que é que lhe digo?” O grupo de instruções de programa chamado *dummy instructions* ou *do-nothing-instructions* é a prova de que o computador não sabe verificar o absurdo de uma alternativa, *formalmente* lógica. Como ao jardineiro, é preciso que lhe digamos: “Manuel! Não faça nada.” (7)

O computador não será capaz de “decisões” não programadas, mas é um nobre animal. Com hábil equitador (no caso o analista) “faz coisas que ninguém acredita que faz” como se canta na marchinha cívico-carnavalesca com que as estações oficiais de rádio comemoram o Sesquicentenário: imprime folhas de pagamento depois de as calcular, dá-nos a indexação das obras de Santo Tomás de Aquino, produz comprovantes de conta bancária, descobre temas musicais do Renascimento, calcula índices do custo de vida, desenha os perfis dos ferros da armadura de uma viga de concreto. E ainda faz versos, escreve romances policiais, compõe *slogans* publicitários e descobre afinidades entre noivos em potencial... (8) até horóscopos é capaz de tirar. Mas tudo depende do programador e do alfabetizador que o ensina a reter relações e correlações. O especialista em arquivos não poderá, pois, ignorar as habilidades de um mecanismo que se tornou indispensável no tratamento de grandes massas de dados, para que delas se possa extrair a informação.

2. Anatomia e Fisiologia do Computador

Ferramenta e manejo — Poderemos distinguir em um sistema-computador um conjunto de partes mecânicas, elétricas e eletrônicas e um conjunto de instruções para manutenção e operação. Estas instruções, ou programas, entram no computador em linguagem capaz de ser compreendida pela máquina e que, por isto mesmo, se chamam “linguagem de

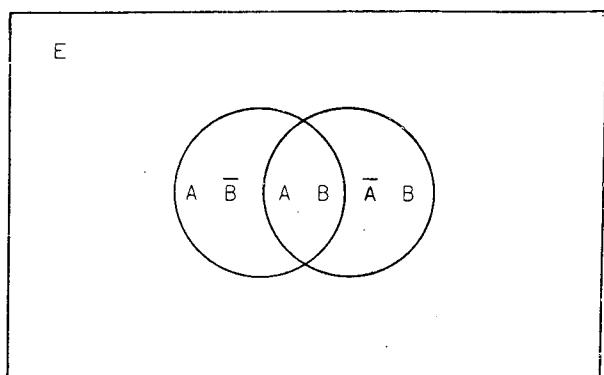
máquina". (º) Alguns programas grandemente aperfeiçoados são fornecidos pelos fabricantes e servem como programa padrão para o desempenho de determinadas tarefas, às vezes muito complexas, como é o caso dos programas operativos. Mudam-se apenas os parâmetros peculiares a cada problema. Outros programas são escritos pelo programador, diretamente, no centro de processamento. O conjunto material do computador se chama *hardware*, que em inglês significa ferragem, armamento ou equipamento de combate. O conjunto de programas e instruções tomou, no jargão de computador, o nome de *software*, modo humorístico de sublinhar a oposição a *hardware*. São termos ainda novos, que Webster ainda não registra na edição de 1968, e que entraram, sem tradução no glossário de informática de todos os países ocidentais. São termos que correspondem, mais ou menos, ao que chamamos de *ferramenta* e *manejo*.

O *hardware* de um computador se compõe, fundamentalmente, de uma unidade de entrada (UE), de um dispositivo de processamento, e de uma unidade de saída (US). A UE alimenta o computador com dados e programas em linguagem de máquina. O dispositivo de processamento se chama processador central ou unidade central de processamento (UCP): efetua operações aritméticas ou lógicas, segundo o programa. A UCP se compõe de uma unidade de armazenamento de dados e programas (memória central), uma unidade de cálculo, capaz de efetuar aplicações aritméticas ou lógicas sobre os dados, e uma unidade de controle e comando, que dirige todas as operações da UCP, verificando-as, diretamente ou em *feedback*, e ainda controla as periféricas. A US restitui a informação elaborada em formato comprehensível, impresso, gráfico, som, microfilme, cópia xerográfica ou imagem em vídeo. A US registra também os resultados intermediários ou finais, em código, em outras memórias de reserva, para processamento subsequente.

A figura 1 mostra a estrutura de um sistema-computador. Hoje este sistema tem grande flexibilidade e a configuração básica pode ser modificada e ampliada à vontade, com a adição de novas unidades periféricas. Nas primeiras gerações de computadores, a unidade central de processamento e as periféricas eram projetadas em conjunto, como sistema único. O controle das periféricas e a transferência de dados entre elas e a UCP eram feitos pela unidade de controle da UCP. Atualmente as periféricas são desenhadas para operar com qualquer UCP dentro, é claro, das características de cada família de computadores. Os dados são transferidos por um elo, que se chama *interface*, e o controle é feito pela própria periférica. Assim, o número de periféricas extras dependerá apenas do número de canais de *interface* na UCP.

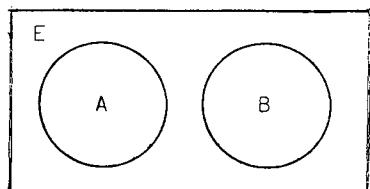
A exposição de funcionamento (fisiologia) do computador nos levaria além dos limites convenientes desta comunicação. Faremos apenas uma observação: os elementos da memória central da UCP são núcleos de ferrita magnetizáveis, em forma de anel ou toro, que recebem, através do programa, uma carga positiva ou negativa. Isto permite uma velocidade

FIGURA 3

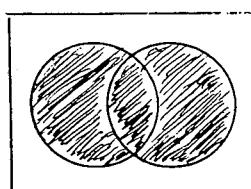


CONJUNTOS COM INTERSEÇÃO

A NOTAÇÃO \bar{X} SIGNIFICA: AUSÊNCIA DO ATRIBUTO X

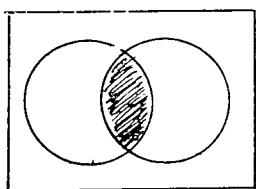


CONJUNTOS DISJUNTOS



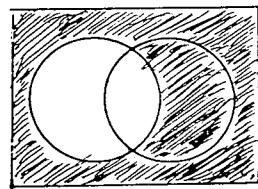
UNIÃO

$A + B$



INTERSEÇÃO

$A \cap B$



COMPLEMENTO

$C(A)$

"E" É O CONJUNTO TOTAL DE DADOS PESQUISADOS

dade vertiginosa de operação, que chega à ordem de grandeza do milionésimo de segundo, para cada operação elementar. Mas, a representação de sinais pela polaridade da carga obriga ao uso de um sistema de codificação binária (ou derivado da binária). O cálculo é feito, pois, em sistema binário (em vez dos dez símbolos do sistema decimal, apenas dois: 0 e 1). Operações em configurações de conjuntos lineares de “zeros” e “uns” constituem a base aritmética ou lógica da recuperação e processamento de dados. O nome computador — recordação das máquinas convencionais de calcular — não exprime toda a gama de tarefas que a máquina pode desempenhar. Razão tiveram os franceses ao chamá-la de “ordenador”.⁽¹⁰⁾ O problema da recuperação de informações, de classificação ou de ordenação, por exemplo, é diferente dos problemas resolvidos por operações aritméticas. A figura 1 nos dá a representação estática (estrutura ou configuração) do sistema-computador, e a figura 2, que mostra o esquema dinâmico, nos dá uma idéia do funcionamento do computador; parece-nos bastante clara e dispensa explicações.

3. Terminologia Informática e Arquivologia

Recordemos algumas noções elementares, do modo mais breve possível: “Comunicação” é um termo geral, designando um sistema que possibilita a interação entre um emissor e um receptor, visando a tornar comum o sentido semântico, operacional ou semântico-operacional de uma mensagem.⁽¹¹⁾ “Informação” no sentido mais amplo, pode definir-se como “efeito da mensagem sobre o receptor”, melhor dizendo, o efeito não seria a “informação” mas a “medida da quantidade de informação recebida”. Se a mensagem é o aspecto *formal* da comunicação, a informação é o aspecto *operacional*. Como resposta a uma indagação do usuário ou como instrução para agir, a informação resulta da elaboração de dados elementares. Fique bem claro que, para uma indagação ou uma instrução subsequente, o que era informação anterior se pode transformar em “dado” ou elemento de informação, para uma informação de segunda ordem, mais complexa.

Os criadores da teoria matemática da comunicação (que foram físicos e engenheiros de comunicação) não levaram em conta o sentido semântico da informação. A única cousa a que visavam era a fidelidade, a integridade da mensagem, em relação a seus sinais, e a eliminação da distorção e do ruído.⁽¹²⁾ Até certo ponto esta noção vale para o interior da UCP, durante o processamento. Não é preciso saber-se “o que significa o que está sendo feito” mas apenas “como está sendo feito”. Mas a preparação dos dados de entrada terá de levar em conta o significado da informação final e o formato de saída terá de oferecer a informação ao usuário, de modo compreensível.

Para que os dados possam ser tratados é preciso que estejam presentes. Quem estiver fazendo uma pesquisa referente, por exemplo, ao comércio brasileiro no tempo dos vice-reis, usará duas espécies de operações:

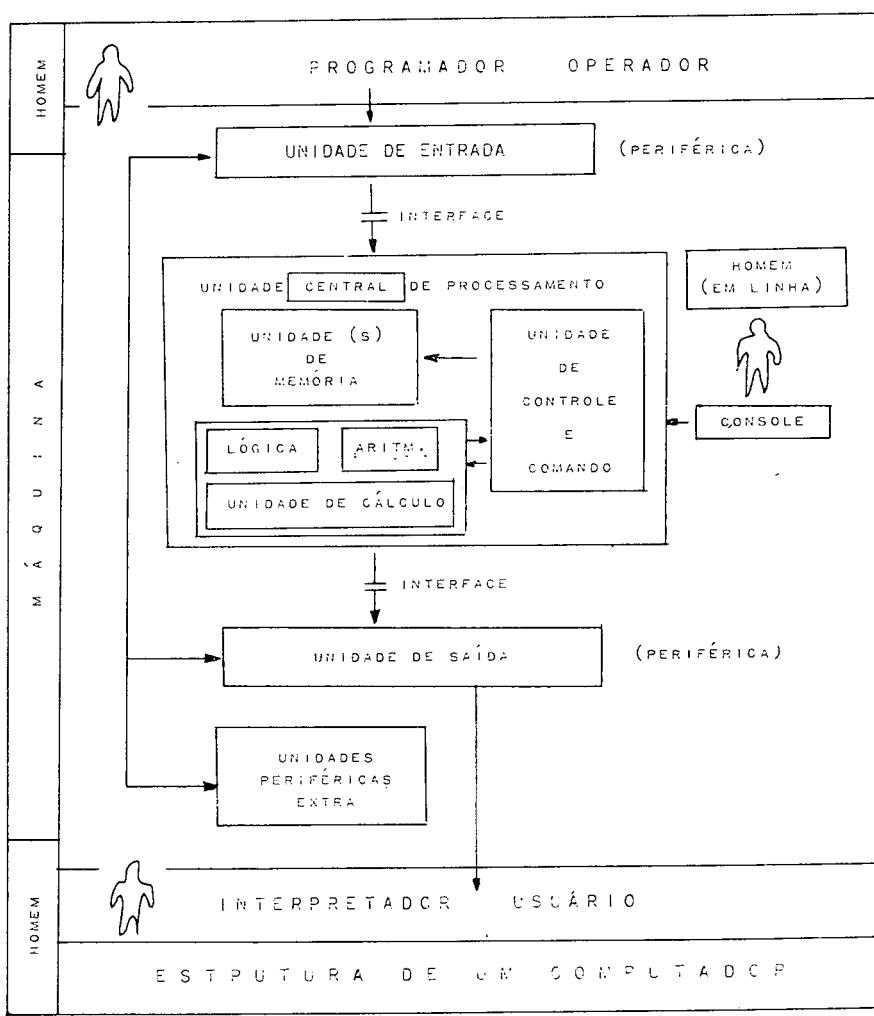


FIGURA 1

a) recuperar todos os dados pertinentes nos maços, livros, cartas, certidões e demais documentos do arquivo; b) elaborar suas notas combinando os dados levantados segundo um “modelo” ou “projeto” que já traz no pensamento. O conteúdo deste segundo “arquivo”, que está na memória do pesquisador, é composto de “lugares” para os dados que espera encontrar nos documentos que estão à sua disposição, e de um “procedimento” para a combinação desses dados. Em uma palavra, um “programa” que o vai guiar na pesquisa. É exatamente o que se passa no computador. Há dois tipos de armazenamento: o da UCP, que é de acesso rápido e direto e o das memórias periféricas, cujo acesso, direto ou seqüencial, exige um procedimento de recuperação. O armazenamento de “dados” e “programas”, na memória central da UCP, corresponde ao conjunto de dados que o pesquisador levantou do arquivo e de procedimentos usados para os combinar, que estavam em sua memória e correspondem a dados e programas que vão para a UCP. Os dados exteriores — no caso o conjunto de dados do arquivo — correspondem aos dados armazenados nas memórias auxiliares,¹³ que podem servir a mais de um programa, como os documentos do arquivo podem servir a mais de um pesquisador. Estes dados, como se passa com as coleções convencionais, podem ou não constituir um arquivo.

Seria útil, neste ponto, a revisão de alguns termos do glossário de Informática, para compará-los com as definições correspondentes do glossário da Arquivística. *Arquivo*, como vimos, é termo polissêmico. Em inglês encontramos três termos para as nossas conotações, *archive*, *store* e *file*, e ainda *memory*, privativo do glossário de computadores. *Archive*, como substantivo (desde 1645 só se usa no plural) significando o local onde se guardam os documentos públicos ou a própria coleção de documentos, mas não o fichário que os contém. Como verbo tem a significação de “arquivar”. O termo *archive* não existe no glossário de Informática. *Store* significa, na linguagem corrente, armazém ou depósito (ou ainda loja de departamentos, no plural). Em Informática é “qualquer dispositivo ou meio (suporte) capaz de registrar informação por certo prazo, permitindo recuperação e uso, quando necessário; é sinônimo de *storage*. Em português se traduz, quase sempre, por memória (de computador) e, raramente, por armazenamento. *File* é o recipiente para os documentos ou a série ordenada dos documentos; como verbo, significa arranjar documentos em certa ordem para preservação e referência.¹⁴ Em Informática, *file* significa “coleção organizada de registros”, cuja relação, ou correlação, depende da finalidade do arquivo e/ou dos procedimentos de inscrição e recuperação. *Memory* que como dissemos é privativo do glossário de computadores, no sentido de armazenamento, no senso estrito indica a memória da UCP e/ou as locações que possam ser endereçadas ou atingidas, em acesso direto, pela unidade de controle da UCP. Às vezes, em inglês, os armazenamentos de acesso direto (discos, por exemplo) são chamados de *memory*. Em português usa-se indistintamente “memória” para qualquer tipo de armazenamento. Por isso, será sempre bom explicitar “memória central”, quando se tratar da UCP.

4. Memória e Arquivo no Computador

Não teria sentido, nesta breve comunicação, estabelecer a diferença entre computador digital e computador analógico.¹⁵ Diremos apenas que no analógico não há, a rigor, armazenamento e memória, há programa. Vamos pois ao digital. Na configuração de um sistema-computador, há vários dispositivos de memória usados tanto para dados como para programas. Podem ser divididos em duas grandes categorias: a) as memórias centrais; e b) as memórias auxiliares ou de massa. Na memória central – de que já falamos – estão os dados e programas necessários à execução de uma tarefa: como os dados levantados pelo pesquisador a que nos referimos há pouco e o seu programa de trabalho. As memórias auxiliares, que se destinam à guarda dos dados e, de certa maneira, correspondem ao que poderíamos chamar de “arquivos” do sistema-computador, podem armazenar os dados em uma enorme variedade de suportes: cartões perfurados, fita de papel perfurada, documentos com caracteres que permitem reconhecimento ótico ou magnético pelas UE, por exemplo, que constituem uma família de acesso menos rápido; a segunda família é constituída pelos suportes magnéticos: fitas, discos, tambores e cartões magnéticos. Podemos ter uma idéia de como funcionam estes suportes olhando a cassete ou o rolo de um gravador de música. A cabeça de gravação/leitura transforma respectivamente as ondas sonoras em *sinais* na película de partículas magnetizáveis e restitui os sinais em forma de ondas sonoras, que vão fazer vibrar os diafragmas do alto-falante. Em vez de fixar na fita (ou no disco, tambor ou cartão, cujo princípio é o mesmo: partículas magnetizáveis que dão gravação ou restituição em trilhas, correspondentes aos sulcos ou às faixas do disco ou da fita usados para gravação musical), em vez de fixar na fita, dizemos, os sinais correspondentes à onda composta do som, gravam-se “sinais” de código, que reproduzem algarismos, letras ou símbolos, formando os registros dos dados necessários à tarefa do sistema-computador e/ou os procedimentos de processamento. As memórias mais empregadas são: a fita e o disco.

A diferença fundamental entre fita e disco está no acesso. Na fita magnética os dados terão de ser gravados em seqüência, à proporção que ela se desenrola no *tape-deck*; qualquer operação de processamento (recuperação ou elaboração) terá de ser feita examinando os registros um a um, na ordem em que foram gravados, até chegar ao endereço ou ao rótulo desejado. O disco magnético, que, como a fita, é recoberto de uma camada sensível (e dos dois lados), permite a gravação ou a leitura em uma área de procura pré-determinada, de acesso direto pelas cabeças de gravação/leitura, montadas em um dispositivo rígido, em forma de pente, uma para cada superfície do disco, capazes de movimento transversal no sentido dos raios do disco. Os discos são arranjados em baterias, e a gravação é feita em trilhas paralelas, não em sulcos em espiral como nos discos de música. As trilhas são divididas em seções e a mesma seção, em uma bateria de discos, forma um setor ou bloco. Cada bloco tem uma capacidade fixa de caracteres, o que significa adensamento crescente da

periferia para o centro; há uma separação ou uma etiqueta, entre os blocos, para distingui-los. Como a finalidade dos discos é dar acesso direto, os dados são sempre organizados de modo a tornar menor possível o tempo de acesso. Como as cabeças estão ligadas rigidamente o “cilindro”, formado pelo conjunto das trilhas correspondentes de todos os discos da bateria, pode ser considerado como uma área de procura, com o respectivo endereço. Indexadas todas as áreas de procura, as cabeças podem ser dirigidas ao endereço desejado, recuperando todos os dados pertinentes em *acesso direto*. Este acesso é também chamado de aleatório ou randômico, em nosso glossário de Informática.¹⁶

Já vimos a diferença entre *store* e *file*. *File*, em um sistema-computador, corresponde rigorosamente a uma das conotações de arquivo de nosso glossário de Arquivística: é uma “coleção organizada de registros” (sistema-computador) ou de “documentos (Arquivo). Há também relação estreita entre registro e documento. Em terminologia de computador documento é “qualquer formulário, ficha ou comprovante contendo os detalhes de alguma transação”. Entende-se por transação “qualquer ocorrência ou acontecimento que exija a geração de um registro para processamento em um sistema-computador”. Como se vê, há uma correspondência estreita entre transação e documento. É só substituir processamento por comprovação.¹⁷

5. Arquivo de Computador e Arquivo Convencional

A diferença maior que se poderia estabelecer entre o “arquivo” do computador e o “arquivo”, no sentido clássico, está na permanência dos dados. Ao revés dos arquivos convencionais, os do sistema-computador são sempre temporários e sujeitos a permanente atualização. O que não quer dizer que não se possa transformar um arquivo de computador (que é sempre registrado em código) em um arquivo permanente em linguagem comum. É só transformar o conteúdo dos suportes de dados do sistema-computador em cópias permanentes (*hard copy*), sejam escritas em linguagem corrente, sejam em gráficos, que também são documentos de arquivo perfeitamente válidas. Além das impressoras de linha e das plotadoras, já existem (há pouco mais de um decênio) duas novidades de saída que facilitam bastante a conversão; são elas: a Saída de Computador em microfilme, SCM (em inglês Computer Output Microfilm, COM) e a Impressora em Xerox, IX (em inglês Xerox Printer, XP), que é uma impressora de página, na qual a configuração de caracteres é feita para a página inteira, antes da impressão em que é usado o processo de xerografia.

Tanto os exemplares permanentes produzidos pelas impressoras de linha, como o microfilme de saída de uma SCM, ou a cópia xerográfica de uma IX podem tornar-se documentos de um arquivo clássico e/ou de um arquivo em microfilme.

Um arquivo magnético não é o mais adequado para a guarda de registros permanentes. Mas, como vimos, é muito fácil copiá-lo em forma convencional. É fácil imaginar quão fecunda será a ligação de um sistema-computador com um arquivo moderno, principalmente no caso de uma instituição ou empresa em que fatos e dados têm de ser constantemente analisados, combinados, quantificados, transformados e sumariados.

6. Lógica do Processamento

O processamento se faz mediante operações lógicas e/ou aritméticas sobre os dados. Como dissemos, o cálculo aritmético não é a única tarefa que pode ser desempenhada por um computador. Embora a maioria das aplicações no sistema-computador (comércio, indústria, administração, pesquisa científica) tenha uma parte preponderante de cálculo, não há, fora das quatro operações elementares, nenhum passo do processamento que não implique aspectos lógicos. As próprias operações aritméticas são redutíveis a operações lógicas sobre conjuntos. O computador efetua apenas *uma* operação aritmética, a soma, e uma operação lógica, a identificação. Através da complementação pode fazer uma operação equivalente à soma de quantidade negativa, que é subtração. Como a multiplicação é apenas uma soma iterada ($3 \times 5 = 5 + 5 + 5$) e a divisão não é mais que uma subtração iterada ($12 : 4 = 12 - 4 = 8 - 4 = 4 - 4 = 0$, no caso da divisão exata e $15 : 4 = 15 - 4 = 11 - 4 = 7 - 4 = 3 < 4 = \text{resto}$) o computador pode efetuar as quatro operações e mais potenciação e radiciação, ainda iterando. Há logaritmos de programação que reduzem a iteração, mas fundamentalmente é isso que acontece. A soma algébrica é uma operação lógica (união de conjuntos mutuamente exclusivos) e a subtração se baseia no complemento de um conjunto. O conjunto de operações lógicas, nas quais estão incluídas soma e complemento, também são o fundamento do tratamento qualitativo da informação, tanto nos registros do arquivo do computador como nos documentos do arquivo convencional. Por isso, não será demais acenar para as operações lógicas básicas. Há, preliminarmente, um procedimento fundamental: a identificação de um dado para sua posterior aceitação ou rejeição (SIM e NÃO). As operações sobre atributos, na lógica de classes são três: a soma lógica ou reunião (OU), o produto lógico ou interseção (E) e o complemento lógico ($\neg A$ ou A). Existem ainda seis operações de lógica de relações, também aplicáveis ao computador.

O quadro 1 nos mostra os dois procedimentos básicos e as três operações da lógica de classes, com a definição e a notação. A figura 3 nos mostra, em diagramas de Euler-Venn, as operações de lógica de classes, esquematizadas.

No tratamento qualitativo dos dados — operação comum aos arquivos convencionais e aos do computador — se fazem aplicações da álgebra de Boole, que é álgebra do SIM e do NÃO. Mas não caberia maiores explana-

ções de um assunto que só se liga tangencialmente ao tema desta comunicação.¹⁸

O quadro 1 e a figura 3 são auto-explicativos: qualquer operação de tratamento de dados terá de fazer-se segundo uma das operações lógicas acima indicadas.

Q U A D R O 1

PROCEDIMENTO BÁSICO

Identificação do dado	aceitação	sim	registra
	rejeição	não	não registra

OPERAÇÕES DA LÓGICA DE CLASSES

(para dois conjuntos)

SOMA LÓGICA ou REUNIÃO	Chama-se reunião de dois conjuntos A e B o conjunto R dos elementos que pertencem a A ou a B.	$A + B$ $A \cup B$
------------------------------	---	-----------------------

Produto LÓGICO ou INTERSEÇÃO	Chama-se interseção de dois conjuntos A e B o conjunto I dos elementos que pertencem a A e a B.	$A \cdot B$ $A \cap B$
------------------------------------	---	---------------------------

COMPLEMENTO LÓGICO	Chama-se complemento de A em relação a E o conjunto de elementos de E que não pertencem a A.	$C(A)$ A'
-----------------------	--	----------------

NOTA — Quando os dois conjuntos são mutuamente exclusivos, a soma lógica é igual à soma algébrica e o produto lógico é o conjunto vazio \emptyset .

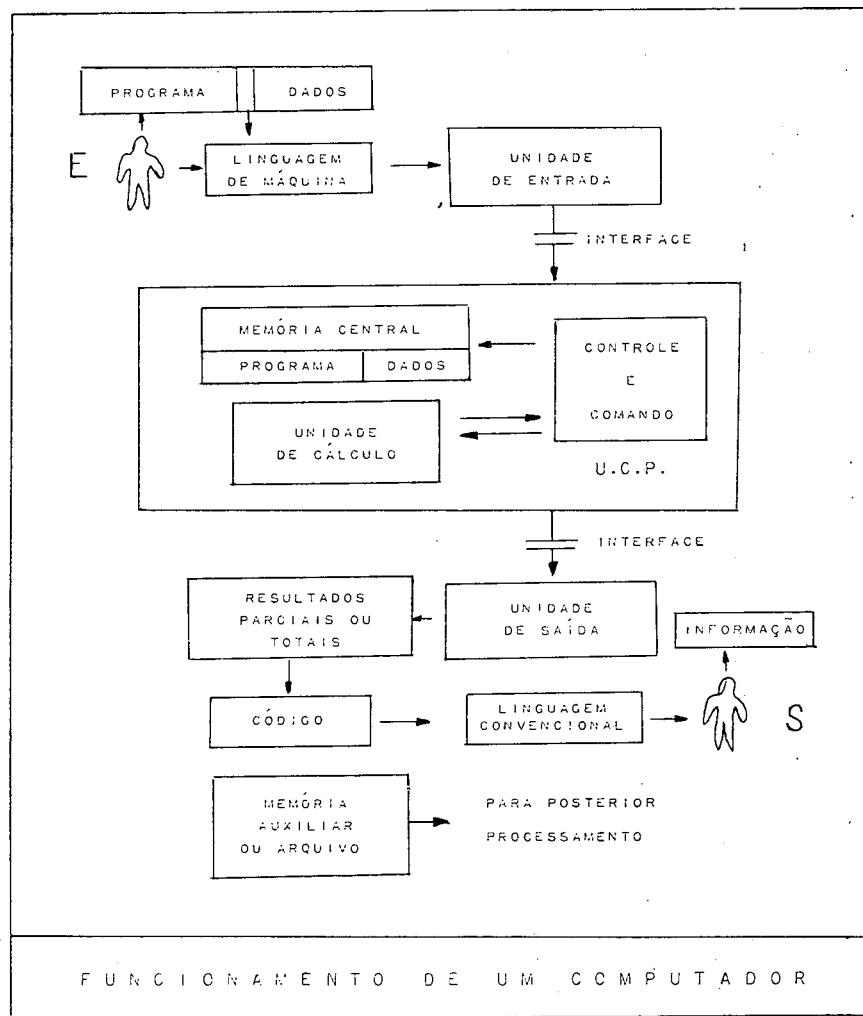


FIGURA 2

7. Tratamento de Dados em Computadores e Arquivos

A rigor, qualquer suporte de informação pode ser usado para a constituição de um arquivo em um sistema-computador. Preferem-se as fitas e os discos magnéticos pela velocidade de acesso, que diminui o tempo de processamento. O que não quer dizer que não haja arquivos de computador em cartão perfurado, fita perfurada ou outro meio de acesso mais lento. Examinaremos o tipo de recuperação nos dois suportes mais usados: fita e disco.

A diferença fundamental está na forma de acesso. A fita terá de ser desenrolada até atingirmos o registro desejado. No disco podem ser fixadas áreas de procura, com endereço, cujo acesso é imediato. O acesso que exige o exame de todos os registros anteriores chama-se “acesso seqüencial”; o que se faz indo, imediatamente, ao registro desejado se chama “acesso direto” (usa-se muito dizer acesso aleatório ou randômico, apesar da incorreção do termo).

Nos arquivos de acesso direto pode estabelecer-se uma sucessão de registros de acordo com determinado seqüenciamento de chaves. Na fita magnética todos os registros, seqüenciados ou não, terão de ser examinados, até chegar se ao endereço ou ao rótulo da série, ou do registro individual, desejados. No disco só se examinam as seqüências ou os registros ligados a chaves predeterminadas. Há, por isso, quem chame de *serial* a entrada de itens sem ordem determinada, reservando *seqüencial* para o arranjo que acabamos de ver. Mas a distinção só vale para suportes de acesso direto. Na fita, *seqüencial* e *serial* são sinônimos.

Há, pois, critérios gerais que poderíamos resumir em três termos: ADJACÊNCIA, ENDEREÇO e COINCIDÊNCIA. Quando todos os registros correspondentes a uma determinada chave estão no mesmo “vaso”¹⁹ ou em vasos contíguos, uma vez chegado o indicador da série, são processados todos os registros adjacentes. O critério é a adjacência. Quando os registros estão dispersos no suporte, será necessária uma indicação de sua posição, que é dada pelo endereço, que permite a busca, seqüencial ou diretamente, conforme o tipo de suporte. Finalmente, se os registros estivessem dispostos serialmente, sem nenhum endereço, teriam de examinar-se todos os registros, até encontrar os que correspondessem a determinada qualificação. Neste caso — que é sempre um caso de passagem — aplicar-se-ia um programa de classificação, transferindo os dados ordenados para um arquivo atualizado, com os dados da mesma família postos em série contínua ou endereçados.

Façamos uma comparação com os documentos convencionais. Um livro não indexado ou uma coleção de cartas ou faturas, na ordem de chegada, correspondem a uma fita na qual os registros foram gravados serialmente. Nos dois casos será preciso uma triagem e uma classificação, formando um novo arquivo atualizado, no sentido da ordem. O livro indexado, ou as cartas já classificadas mas ainda não ordenadas, corres-

pondem à busca através do endereço. Finalmente, se dispusermos os registros, separadamente, em pastas ou maços, contendo os dados ordenadamente, por assunto, isto é adjacentes, teremos o outro caso. Sublinhamos essa coincidência formal nos procedimentos de pesquisa e recuperação. Haveria muito ainda que dizer sobre as correspondências entre um arquivo de computador e um arquivo convencional.

8. Necessidade de um Trabalho Conjunto

Acenamos, muito resumidamente, com alguns pontos que mostram a necessidade de colaboração mais íntima entre todos os que trabalham no imenso e difícil campo da Informação. O arquivo — e não só pelos leigos — continua a ser considerado como algo de morto, quase que um mal necessário para que sejam satisfeitas exigências legais ou resguardadas tradições e lembranças históricas. Com a explosão da informação, a importância fundamental do ARQUIVO vai sendo reconhecida em todas as províncias da atividade humana.

Os antigos já sentiam toda a riqueza de conotações da palavra ARQUIVO. Há mais de cento e oitenta anos, o dicionarista Antonio de Moraes e Silva apresentava dois exemplos interessantes da extensão do significado de “arquivo”. “Sua memória era um arquivo de vastíssimas erudições”. Exatamente o que dizíamos, ao comparar a memória humana do pesquisador com a memória eletrônica da UCP. E logo este outro: “tudo isso estava guardado nos arquivos da graça divina”.²⁰ Até Deus tem seus arquivos e seus livros de registro, como nos diz em formosa imagem Tommaso di Celano, um dos maiores poetas da Itália medieval:

Liber scriptus proferetur
In quo totum continetur
Unde mundus judicetur.²¹

Não será, pois, demasia que todos os que se encontram no campo da Informação, profissionais ou usuários, procurem compreender as linhas gerais dos procedimentos, metodologia e teoria geral dos arquivos. Para os especialistas isto é uma necessidade inadiável: é preciso que todos os ramos da Teoria e da Prática de Arquivos se unam à Biblioteconomia, Bibliologia, Museologia, Documentação, Análise de Sistemas, Informática, Administração e Direito, trabalhando, todos juntos, para o estabelecimento de uma tecnologia avançada, de uma normalização de glossários e de um sistema de normas e regulamentos, para que o trabalho da guarda e da recuperação da Informação, para posterior disseminação, seja um trabalho eficiente e eficaz. Hoje, em Administração Pública ou de Empresas já não se encara o arquivo como mero depósito daquilo que passou, senão como um meio seguro, imediato, flexível e eficiente para o estabelecimento de programas e a tomada de decisões fundamentadas. O arquivo é instrumento ao mesmo tempo retrospectivo, atual e prospectivo. É preciso que a nova abordagem de arquivo se estenda a todos os campos da atividade humana.

BIBLIOGRAFIA

1. BERKELEY, Edmund C. & LOVETT, Linda L. — *Glossary of terms in computers and data processing*. Newtonville, Mass. Berkeley enterprises, 1960.
2. BOWLES, Edmund Addison. ed. — *Computers in humanistic research; readings and perspectives*. Englewood Cliffs, N.Y., Prentice Hall /c1967/
3. CHANDOR, Anthony, GRAHAM, John & WILLIAMSON, Robin — *A dictionary of computers*. /Harmondsworth, Engl./ Penguin /1970/
4. ——— — *Practical systems analysis*. London, Hart Davis, 1969.
5. ESTADOS UNIDOS — *Glossary of automatic data processing*. Washington, D.C., Govt. Print Office, 1962.
6. GOLDSTEIN, Jaime — *Sobre a informação jurídica*. Rio de Janeiro, PUC, 1972 (pré-impressão).
7. LUCENA, Carlos Pereira de — *Introdução às estruturas de informação*. Rio de Janeiro, Livro Técnico, 1970.
8. MACK ADAMS, J. & MOON, R. — *An introduction to computer science*. New York, Scott Foresman, 1970.
9. MIRANDA NETTO A.G. de — Informação, documentação, cibernética. In: *Informática*. Rio de Janeiro, IBBD, 1969.
10. ——— — *Terminologia científica e técnica, barreira à comunicação e à automação*. Rio de Janeiro, IBBD, 1969.
11. MOLES, A. — *Theorie de l'information et perception esthétique*. Paris, Flammarion, 1958.
12. ——— — Archives. In: *La communication*. Paris, Danoel, 1971.
13. OLIVEIRA, Maria de Lourdes Claro de & ROSA, José Lázaro de Souza — *Teoria e prática da microfilmagem*. Rio de Janeiro, Fundação Getúlio Vargas, Instituto de Documentação, 1970.
14. PAFS, Marilena Leite — *O papel da arquivística na documentação*. Rio de Janeiro, Fundação Getúlio Vargas, Instituto de Documentação, 1971.
15. POULAIN, Pierre — *Éléments fondamentaux de l'informatique*. Paris, Dunod, 1969.
16. POURCELL, R. — De la machine de Pascal aux ordinateurs de demain. *Informations sociales*, Paris, julho-setembro, 1968.
17. SAMMET, J. — *Programming languages, history and fundamentals*. New Jersey, Prentice-Hall, 1969.
18. SHANNON, C.E. & WEAVER, W. — *The mathematical theory of communication*. Urbana, Un. of Illinois Press, 1949.
19. SILVA, Benedicto — *Origem e evolução dos descritores*. Rio de Janeiro, Fundação Getúlio Vargas, Serviço de Publicações, 1972.
20. TATON, R. & FLAD, J. P. — *Le calcul mécanique*. Paris, PUF, 1949.
21. VASCONCELLOS, Augusto de — *Computadores eletrônicos digitais*. Rio de Janeiro, Liv Kosmos /1971/
22. ——— — *Programação de computadores eletrônicos; conceitos básicos*. Rio de Janeiro, Liv. Kosmos /1971/
23. VERDE, Raul — *Análise e programação de computadores*. Lisboa /Liv. Quadrante, 1971/

INTERVENÇÕES

De RAUL LIMA

É um encanto ouvir Miranda Neto, a quem o Congresso deve o seu maior show, sem prejuízo da importância dos ensinamentos. A Mesa-Redonda da Conferência Internacional de Arquivos, no ano passado, em Bonn, teve como um de seus dois temas o Arquivo e a Informática. E os pronunciamentos dos grandes Arquivólogos e as conclusões foram de extrema prudência em relação ao uso de computação em Arquivos de custódia. Pelo que ouvi, parece que você também não recomenda. Quer ensinar-me alguma coisa a respeito?

R) Em matéria de recuperação o computador terá de ser empregado com grande cautela. Fundamentalmente o computador faz duas coisas: *soma* e *separa*. Com a soma chega a todas as operações aritméticas. Com a separação pode prestar serviços notáveis na recuperação documental. A aplicação do computador à pesquisa legal é das mais interessantes. Quanto aos arquivos de custódia creio que se deve examinar o custo do computador (ou do terminal) e da programação. É quase certo que o trabalho normal é mais barato e de igual eficiência. Estamos por demais encantados com a tecnologia. De mim, ainda sou fiel ao velho dito latino "Ó livros amigos, é doce viver entre vós e entre vós morrer!" Veria com extremo desgosto o meu Balzac em uma tela de terminal. Leia *Computers and Common Sense* de Mortimer Taube. É um bom antídoto contra a epidemia de informática que anda por aí.

De MARIA LUCIA FANTIN

Poderia abordar novamente o "discutido" e mal aplicado conceito de informática?

R) Informática, segundo o criador da palavra, Dreyfues, é "a ciência (eu prefiro dizer tecnologia) do tratamento automático e racional de informação, considerada como suporte de conhecimento e de comunicações". Está em geral ligado ao computador e o norte-americano prefeira usar "computer science" (ciência do computador).

De GUY DE HOLLANDA

Parecer haver um estrangulamento na passagem dos dados de documentos microfilmados para a alimentação do computador. O reconhecimento ótico ou magnético ainda não permite fazer diretamente esse trabalho?

R) Um dos grandes problemas do computador é o estrangulamento devido à diferença de velocidade entre o homem e a unidade central de processamento. O mesmo se dá em relação às periféricas lentas. A programação demora bastante e as técnicas de multiprocessamento procuram obviar esse inconveniente. O reconhecimento ótico (ou magnético) ainda está na infância, mas os êxitos obtidos no reconhecimento de formas (pattern recognition) permitem esperar bastante para o futuro.

De NILZA T. SOARES

O computador é fabuloso, mas ele vai atrapalhar o nosso tradicional "jeitinho". Quando a coisa é programada, nada se pode fazer, não se pode alterar prazos, por exemplo: o que o senhor acha?

R) Creio que o computador jamais substituirá o homem, enquanto não for capaz de autoprogramar-se, o que me parece, por enquanto, impossível. Quanto ao tradicional "jeitinho", o programador pode encontrar meios de o fazer. Já li exemplo de fraudes imensas que passam despercebidas, dada a habilidade do programador. O problema de auditoria de processamento é, por isso, um dos mais sérios.

MOÇÕES, SUGESTÕES, CONGRATULAÇÕES E PERGUNTAS NÃO RESPONDIDAS

De ERNESTINO FIGUEIRA

O Senhor tem obras publicadas? Quantas? Gostaria de comprar ou receber as referentes à parte arquivística.

De ANA MARIA CABRAL

Nunca ouvi alguém que comunicasse tanto e prendesse a minha atenção. O Senhor é genial e muito simpático.

De ALUNAS DE BIBLIOTECONOMIA DA UFF

Encantadas com a palestra.

De UMBERTO PEREGRINO

Estão presentes, especialmente trazidos à sessão de hoje, alunos do Curso de Comunicação e Editoração da UFRJ, em nome dos quais agradeço a sua primorosa exposição, ao mesmo tempo tão válida e saborosa.

De WANDA MENDONÇA ALENCAR

Tomarei o empenho de mandar publicar na imprensa de Manaus sua brilhante palestra. O Brasil precisa de homens de valor como o Dr. Antonio Garcia de Miranda Netto.