

O «Centre National de la Recherche Scientifique» (Paris)

por MARIA DA CONCEIÇÃO OSÓRIO GONÇALVES
Bibliotecária da Biblioteca Nacional de Lisboa

RESUMO: A contribuição do CNRS (Paris) no domínio da informação científica e técnica deve-se sobretudo à publicação do *Bulletin Signalétique* presentemente dividido em cinquenta e cinco secções especializadas. O crescimento contínuo da documentação levou a uma avaliação e reorientação de todo o sistema e métodos de trabalho do Centro. É neste sentido que se desenvolve todo o artigo ao analisar os seguintes pontos: *Estruturas, Catalogação, Indexação e Thesauri, Ficheiros de palavras-chave, Análise bibliográfica, Tratamento automático da informação, Gestão automatizada dos periódicos recebidos, Reprografia, Valor da informação recebida, Custos da informação.*

SUMMARY: The contribution of the CNRS (Paris) in the field of scientific and technical information lies above all in the publication of the *Bulletin Signalétique*, presently divided into fifty five specialised sections.

The constant increase of documents has brought about a revaluation and new orientation of the entire system and operation of the centre.

This is the sense in which the entire article is developed, analysing the following topics:

Structures, Cataloguing, Indexation and Thesauri, Keyword files, Bibliographical analysis, Automatic treatment of data, Automatic cataloguing of all periodical publications received, Reprography, Worth of received informations and their costs.

1. O CNRS tem uma longa história, que data de 1940, encontrando-se neste momento numa fase de reorientação.

Correspondem a sua criação, e ainda a maior parte dos seus sistemas e dos seus métodos, a uma etapa da biblioteconomia e da ciência documental em que predominava o estabelecimento de grandes bibliotecas e de grandes centros de documentação que deveriam responder às necessidades da informação científica e técnica em todos os domínios de estudo.

Os serviços prestados pelo CNRS, no decurso deste período já tão longo, devem-se fundamentalmente à publicação do *Bulletin Signalétique*, presentemente dividido em cinquenta e cinco secções especializadas. A sua publicação, desde uma época em que tais iniciativas eram ainda em muito pequeno número, seria suficiente para justificar, se tal fosse necessário, a existência e o trabalho realizado pelo CNRS.

É evidente que o crescimento contínuo da documentação e a criação de centros de documentação especializados põem problemas de base a um centro deste tipo, levando necessariamente a uma avaliação e reorientação de todo o sistema e métodos de trabalho. Como primeiros resultados: quatro das secções do *Bulletin Signalétique* são já publicadas em comum com os Centros de Documentação do *Bureau des Recherches Géologiques et Minières/Paris, Institut Pasteur/Paris, Institut de Soudure/Paris, Service Central de Documentation du C.E.A./Saclay*; a automatização do sistema, iniciada em 1971, embora ainda numa fase experimental, tornou já possível a elaboração automatizada do *Bulletin Signalétique*, bem como a execução de perfis s.d.i. (selective dissemination of information) individuais e de grupo em alguns domínios. É evidente que todas as pesquisas retrospectivas anteriores a essa data terão de ser feitas manualmente através dos numerosos ficheiros existentes.

2. Estruturas

O sistema possui, como elementos principais, a Redacção do boletim, o Serviço de Informática e respectivo Laboratório, os Serviços Técnicos, a Biblioteca, os Serviços de Reprografia.

Os periódicos entrados (cerca de 13.000 «vivos», dos quais uns 9.000 são indexados) e as obras avulsas são primeiramente submetidos à apreciação da Redacção que julga do seu interesse. Os que não são considerados válidos para publicação são enviados à Biblioteca que por sua vez decide da vantagem de os conservar ou não.

Os documentos analisados são catalogados pelos Serviços Técnicos, que se encarregam igualmente da sua circulação, enviando-os, com os respectivos impressos parcialmente preenchidos: à Redacção para indexação e publicação; a pessoas contratadas, fora do CNRS, para elaboração das respectivas análises bibliográficas; a Organismos que colaboram com o CNRS para indexação e redacção das mesmas análises; aos quatro Organismos que, em colaboração com o CNRS, publicam uma secção do *Bulletin Signalétique*; por fim, para armazenagem, à Biblioteca, que novamente os cataloga.

3. Catalogação

A catalogação põe neste momento alguns problemas. Em primeiro lugar é feita duas vezes, segundo normas diferentes: uma pela Biblioteca, que segue as normas internacionais e francesas, acompanhando a catalogação da Biblioteca Nacional de Paris, outra pelos Serviços Técnicos que preparam os impressos para publicação no boletim e consequente tratamento automático. Esta segunda catalogação é muito simplificada, existindo instruções precisas feitas em função do impresso a preencher. É essencialmente prática, abandonando o estudo teórico da catalogação e utilizando possibilidades próprias de um sistema automatizado (por exemplo: os nomes dos autores compostos por vários elementos dão entrada, na ordem alfabética, por cada um desses elementos, o que elimina a necessidade de conhecer o carácter linguístico de cada país. Se é verdade que tal facto simplifica o trabalho exigido, é-o também que cria uma desfasagem importante em relação ao tratamento não automático dos documentos. É impensável, nestes casos, sobrecarregar os ficheiros com um tão grande número de fichas).

Ainda quanto à existência de duas catalogações, é evidente que tal facto implica uma grande sobrecarga de trabalho. O CNRS aguarda contudo a publicação das regras UNISIST para estabelecer uma coordenação.

4. Indexação e Thesauri. Ficheiros de palavras-chave

A indexação dos documentos feita, quer por especialistas quer por técnicos, é executada segundo critérios variáveis. Os vocabulários utilizados formam, nuns casos, *thesauri*; noutros constam de ficheiros de palavras-chave já com o correspondente ficheiro invertido; noutros ainda não existe ficheiro invertido, mas apenas ficheiros de trabalho. Embora esteja a ser feita a automatização, para elaboração de perfis s.d.i., das secções em que existem já ficheiros invertidos, ou seja, ficheiros numéricos de descritores ou palavras-chave. Dado que estes mesmos ficheiros não estão linguisticamente coordenados entre si, são muito numerosas as repetições de vocábulos dentro do mesmo domínio científico, de acordo com as secções em que está subdividido. Isto implica, evidentemente, que cada secção possua um código próprio, sendo necessário, para a elaboração de perfis, utilizar as fitas magnéticas segundo códigos diversos.

Embora sendo a questão da terminologia uma das mais difíceis de resolver, dada a sua grande variação, inclusive dentro de um determinado domínio, uma situação deste tipo é certamente muito imperfeita do ponto de vista técnico. (*Thesauri* que podem ser pedidos ao CNRS: *Pathologie générale et expérimentale*, 200F; *Sciences pharmacologiques*, 300F; *Géophysique: I — Termes généraux*, 150F; II — *Termes géographiques*, 80F; *Polymères*, 250F; *Vocabulaire des termes rencontrés sous diverses formes en chimie*, 60F).

Há ainda outro aspecto a considerar que implica uma variação de qualidade na indexação feita. Além dos quatro organismos que publicam uma secção do boletim em conjunto com o CNRS, e que na prática se responsabilizam inteiramente por cada uma das secções publicadas, trinta e dois outros colaboram na análise dos documentos, fazendo mesmo alguns a respectiva indexação. A variação de critérios de indexação, que sempre existe, atinge imediatamente uma amplitude muito maior.

5. Análise bibliográfica

As análises bibliográficas, com excepção das que são feitas pelos organismos que colaboram com o CNRS, são elaboradas por especialistas, contratados fora dos quadros do Centro.

6. Tratamento automático da informação

A automatização, a partir de 1971, do tratamento de dados informativos recolhidos, originou todo um novo sistema, o sistema PASCAL (*Programme Appliqué à la Sélection et à la Compilation Automatique de la Littérature*). Os elementos referentes a cada documento são gravados *en vrac*, sem qualquer distinção entre os vários domínios a que se referem. Desta primeira fita são em seguida extraídas outras duas, segundo programas diversos, destinando-se uma à elaboração do *Bulletin Signalétique* e a segunda à execução de perfis s.d.i., nos casos em que já é possível dispor de um ficheiro invertido.

6.1. Bulletin Signalétique

Uma vez por mês são portanto extraídos da primeira fita magnética gravada os elementos relativos a cada secção do boletim. São codificados, são-lhe atribuídas rubricas e são enviados para fotocomposição. A fotocompositora utilizada, que lê fita magnética, permite uma impressão em maiúsculas e minúsculas, utilização de caracteres gregos, símbolos matemáticos, sinais diacríticos. Isso implica, ao nível do programa, uma «recolha rica» dos elementos de escrita.

Além dos dados informativos, cada rubrica contém ainda índices de autores e de assuntos.

6.2. Programas s.d.i.

Uma segunda recolha de elementos serve portanto a difusão selectiva da informação, utilizando de momento um sistema denominado MISTRAL e estando em estudo a utilização de um sistema sueco, VIRA. No sistema MISTRAL os dados são gravados em fita magnética. As perguntas são codificadas em disco. São estas que percorrem os dados registados, pelo que o custo é igual ao número de perguntas. Trata-se de um sistema aconselhável, por exemplo, para uma recolha retrospectiva da informação, mas caro para perguntas frequentes. O sistema VIRA compara simultaneamente um grande número de perguntas, codificadas em fita magnética, com um grande número de dados, igualmente gravados em fita magnética. É rentável apenas para um grande número de perguntas, mas, nesse caso, mais económico do que o anterior. Para além de cem perfis, um perfil fica praticamente gratuito.

Quer um sistema quer outro implicam, como já vimos, a existência de um ficheiro invertido; neste momento tais ficheiros, ou códigos, estão preparados para trinta e duas secções das cinquenta e cinco existentes.

6.3. Hardware

O CNRS utiliza os computadores Siemens 4004/35, IBM 3600/30 e IBM 370/165, em sistema de *part-time*. Para fotocomposição são utilizadas as máquinas DIGISET e LINOTRON 505.

7. Gestão automatizada dos periódicos recebidos

Anterior à automatização do sistema informativo foi o projecto que pretendeu criar, em 1968, um sistema automático de controle dos periódicos recebidos. Em linhas gerais, foram analisados os condicionalismos de chegada dos periódicos — irregularidade de publicação, alteração de periodicidade, alterações de títulos, prazos postais — sendo tidos em conta na programação. O sistema consistia fundamentalmente em indicar as datas prováveis de chegada de um novo número de cada periódico; eram em seguida regularmente fornecidos pelo Laboratório os cartões dos números a serem recebidos num determinado prazo, a fim de ser feito o respectivo controle de entrada ou as reclamações de números em falta.

Este programa foi abandonado, pois a não existência de computador no CNRS obrigava a um movimento contínuo de grande número de cartões, o que se tornava impraticável. Os elementos bibliográficos descritivos dos periódicos continuaram contudo a ser armazenados, tornando possível a elaboração de um catálogo dos periódicos

existentes no CNRS. Este catálogo indica, para cada periódico, apenas a data de início de colecção e não os números existentes. Está todavia a ser testado um programa de actualização que fará a fusão da primeira fita magnética gravada com uma segunda fita magnética contendo os elementos relativos ao estado das colecções. Tais elementos serão regularmente enviados ao Laboratório em cartões perfurados.

Quanto à automatização da gestão dos periódicos, não há, de momento, qualquer novo processo em estudo no CNRS, mas considera-se praticável apenas em tratamento *on-line*.

8. Reprografia

São pedidos ao CNRS grande número de fotocópias, microfichas e microfilmes. Dado que, como vimos, nem todos os documentos analisados são armazenados, o CNRS montou em várias bibliotecas de Paris um serviço próprio de fotocópia. É-lhe assim possível fornecer a quase totalidade dos documentos pedidos. Existem, contudo, alguns cuja análise lhe é enviada por outros Organismos, a que não tem acesso.

9. Valor da informação recebida

As imperfeições evidentes do sistema de informática, motivadas em parte por uma automatização que partiu de uma base não suficientemente preparada, desaconselham os modelos adoptados, dificultam o trabalho dos técnicos, mas não implicam necessariamente uma desvalorização da informação fornecida. De maior importância são já as variações de indexação. Quanto à falta de coordenação dos vocabulários, influi inevitavelmente na recolha da informação, mas pode ser em parte numerada através de um bom trabalho técnico de que o utilizador não se chega a aperceber. Interessaria conhecer como é feita a actualização dos vocabulários e poder avaliar a percentagem das respostas pertinentes através de uma ou duas perguntas feitas simultaneamente a vários sistemas. Não existe qualquer estudo deste tipo feito relativamente ao CNRS.

Seria igualmente necessário um estudo comparativo, com outros Centros de Documentação, em relação a cada domínio tratado, da documentação coberta. O CNRS não dispõe de momento de dados exactos relativos aos vários domínios trabalhados, tendo, contudo, indicado alguns números referentes a Ciências da Terra, Economia e Medicina. Por terem sido calculados por aproximação não os incluímos neste Relatório.

10. Custos da informação

10.1 Fitas magnéticas

Os preços das fitas magnéticas variam, de acordo com as várias secções, entre 3 000 e 15 000F, por ano e por secção. Das vinte secções referentes a Medicina não existem à venda fitas magnéticas das secções 320-Bioquímica e Biofísica e 351-Cancro.

10.1.1 Preços por secção para 1973

101	Information, Documentation	2 800F
110	Informatique, Automatique, R. O., Gestion	5 000F
120	Astronomie, Géophysique	5 900F
140	Electrotechnique	4 000F
145	Electronique	5 900F
150	Physique nucléaire	5 000F

170	Chimie	15 000F
310	à 390, moins 320 et 351-Médecine	51 000F
730	Combustibles, Énergie	6 000F
740	Métaux, Métallurgie	7 000F
745	Soudage	4 000F
780	Polymères	6 000F
880	Génie chimique	5 000F
885	Nuisances	4 000F
920	Sciences de la terre	10 000F

Estes preços são acrescidos do custo dos suportes.

10.2 *Perfis*

Perfis de grupo	250F por ano
Perfis individuais	400F por ano
Perfis confidenciais	a fixar caso por caso

10.3 *Informação retrospectiva* (período de 5 anos)

Por assunto	200F
Por autor	50F
Outras informações	30F

Estes preços são acrescidos de 100F caso se trate de um assunto com um número elevado de referências ou acrescido de investigação complementar de bibliografia de outros países.