
CIÊNCIA DA INFORMAÇÃO: CIÊNCIA DA FORMA*

Richard Bouche
Laboratório de Informática Documentária
Universidade Claude Bernard
Lyon I — França.

RESUMO

A disciplina de ciência da informação e da comunicação aparece como um novo campo científico. O que representam as ciências da

informação nesse quadro? A idéia que se desenvolve é que ciência da informação consiste em dar forma à mensagem em função de um canal específico e um objetivo fixo da comunicação. Este aspecto, aparentemente de ordem técnica, na realidade diz respeito à pesquisa fundamental. Por exemplo, no campo do documento escrito é o problema da conservação da função da referência extralingüística que interessa. Considerações desse mesmo tipo podem ser feitas a propósito de outras formas de informação (imagens fixas, imagens animadas, emissão radiofônica, impressão de um livro etc.). O domínio da ciência da informação aparece então como um vasto campo ainda inexplorado.

1 - INTRODUÇÃO

Na França, há dez anos, foi criado no Ministério de Educação Nacional uma seção de ciências da informação e da comunicação servindo de quadro às atividades de ensino e de pesquisa. Essa criação corresponde ao desenvolvimento de uma nova disciplina, na qual o campo científico se define pouco a pouco. Dessa mesma forma, em outros países, assiste-se a uma evolução de escolas de biblioteconomia e de documentação em direção a uma demoninação contendo a expressão "ciência da informação". Qual é o domínio dessa disciplina, como ela se situa em relação às outras ciências da comunicação (se é o caso de se fazer uma distinção)? No que se segue tentaremos dar uma resposta a essas questões, mostrando que uma definição do campo científico das ciências da informação pode ser feita, que ele é vasto e que ainda permanece em parte inexplorado.

2- A COMUNICAÇÃO

Ainda que sujeito a certas críticas, o esquema de Jakobson é freqüentemente utilizado para descrever os fenômenos da comunicação.

— O *emissor* que, em certo momento, aprovou a necessidade de emitir uma mensagem. Esta pode ser o resultado de uma iniciativa pessoal ou efeito de uma pressão social;

— O *receptor* é o destinatário da mensagem. Ele é mais ou menos identificado. O resultado da comunicação se traduz por uma certa satisfação ou insatisfação que está freqüentemente na origem da pressão social citada no item anterior.

— O *canal*. É por ele que transita a mensagem. Ele se compõe de elementos físicos (papel, filme, vibrações sonoras, ondas eletromagnéticas etc.) e precisa utilizar ferramentas mais ou menos complexas (caneta, microfone, camera, aparelhos de telefone, receptores de televisão etc.). O canal possui características precisas; possui igualmente imperfeições que vão trazer perturbações à mensagem sob forma de distorções ou de ruídos.

Mas, outros dois aspectos podem ser igualmente considerados:

— O *conteúdo da mensagem*. Se é preciso atribuir-lhe um valor, este depende não apenas da mensagem propriamente dita mas também do contexto da comunicação. O emissor vai se esforçar para satisfazer o que supostamente ele acha que o interlocutor espera. Quanto ao receptor, se ele identificou o emissor, a imagem que ele faz do outro vai colorir fortemente o valor da mensagem recebida.

— A *forma da mensagem*. Ela deve corresponder aos objetivos visados e às características do canal escolhido.

* Tradução do original francês por Márcilio de Brito, da Secretaria Especial de Desenvolvimento Industrial do Ministério da Indústria e Comércio.

3 - A INFORMAÇÃO

Experimentemos agora definir a palavra informação. Aqui encontra-se a mesma oposição anterior entre a forma e o fundo.

Como várias palavras com sufixo em /ção/, a palavra informação possui dois sentidos: ela significa ao mesmo tempo o resultado do ato de informar e o próprio ato.

— Se considerarmos o resultado do ato, colocaremos em evidência um valor informativo, que se situa sobre o conteúdo. A informação assume valores subjetivos e fortemente ligados ao contexto econômico e social. Os efeitos e os problemas atribuídos ao canal são ignorados.

— Quanto ao ato propriamente dito, ele se reveste de um aspecto técnico. Que atitude adotar para informar? Em caso limite, como na teoria de Shannon, a informação não é mais que uma sucessão de *bits*, cujo valor informativo está ligado apenas ao caráter mais ou menos improvável de suas aparições em uma mensagem.

Notaremos que, etimologicamente, a palavra "informar" vem do latim "informare", significando "colocar em forma". Esta etimologia parece assim privilegiar a forma sobre o conteúdo; produziu-se, em relação ao sentido atual, um deslizamento do sentido, passando da forma ao conteúdo.

As conseqüências das considerações precedentes nos conduzem a dizer que só existe comunicação a partir do momento em que o receptor recebe a informação que transitou pelo canal. Mas é preciso considerar não apenas o espaço, mas também o tempo. O exemplo clássico citado por A. Moles¹ é o estabelecimento de uma ligação de comunicação entre Madame de Sévigné e um leitor de 1988 através da história por intermédio de um livro.

4-CIÊNCIA DA INFORMAÇÃO-CIÊNCIA DA COMUNICAÇÃO

Partindo dessas observações, daremos nossa definição de informação e de ciência da informação.

A informação é uma forma que circula sobre o canal, do emissor ao receptor. Ela não possui valor em si. Para o emissor ela possui um valor de partida, mas adquire um outro valor mais ou menos diferente do precedente para o receptor quando ele a recebe.

A ciência da informação é, por assim dizer, a ciência do condicionamento a uma forma e da adaptação a um canal.

A ciência da informação é uma parte da ciência da comunicação que diz respeito ao ato completo da comunicação, em particular seus aspectos humano, econômico, jurídico e social.

Esta definição de ciência da informação parece reduzida aos aspectos técnicos; veremos, contudo, que isto representa um campo de interação com outras disciplinas, largamente variadas e mesmo muito antigas e suficientemente vastas e ricas para chegarem às pesquisas de caráter fundamental, podendo assim enriquecer com retorno as outras disciplinas.

Esquema 1-1 (1 emissor- 1 receptor)

Consideremos em primeiro lugar um esquema simples de comunicação dentro do qual o emissor tomou a iniciativa do contato e sabe com qual interlocutor trabalha, mesmo se for a primeira vez que uma tal conexão é estabelecida.

Exemplos de comunicações deste tipo são o correio, a conversação face a face ou telefônica. Esta última é considerada com mais atenção. Na verdade existem dois níveis técnicos:

- um nível elementar, ou nível de base, relativo à ligação física: o telefone, os comutadores, os fios, os "cabos" hertzianos, os satélites etc. Trata-se do suporte da informação, e a pessoa que está implicada neste nível é um físico, químico, engenheiro mecânico, engenheiro elétrico, acústico etc.
- um outro nível mais abstrato, ou nível intermediário, diz respeito à utilização das ferramentas precedentes para que a mensagem transite bem. A comunicação estabelecida possui: um objetivo; trata-se de convencer o interlocutor do bom fundamento de um contrato, de uma colaboração, de um encontro etc. A forma da mensagem deve ser adaptada a este objetivo e ao canal utilizado.

Pode-se pensar em outros esquemas mais complexos, nos quais um emissor se dirige a vários receptores ou, situação inversa, um receptor recebe mensagens de vários emissores.

Esquema 1-n (1 emissor- n receptores)

Trata-se aqui dos meios de comunicações (imprensa, edição, rádio, televisão, cinema), mas também das conferências e, por que não, do teatro, do concerto, da dança.

O suporte técnico é diverso e depende do canal. Técnicos são encarregados de criar os instrumentos de base necessários que podem ser

suficientemente complexos (por exemplo, uma cadeia de televisão com todos os equipamentos de tomadas de imagens, iluminação e decoração, mas também todos os dispositivos de mixagem e trucagens permitindo intervir sobre a imagem).

Encontramos, freqüentemente com uma complexidade maior, o nível técnico intermediário observado no esquema precedente. Ele consistiria em dar forma à mensagem para que ela atinja o usuário indiferenciado, "o público" que pode ser mais ou menos definido.

No caso de uma conferência, a competência em matéria de manipulação de instrumentos (microfone, retroprojeter), ou de retórica, permite ao conferencista passar com sucesso sua mensagem.

Um bom autor saberá escrever um bom livro, mas o papel do editor sob a forma de livro e sua distribuição é igualmente um elemento não negligenciável do sucesso da mensagem.

A difusão de um artigo depende muito da maneira como ele é escrito, mas também de sua forma e de seu lugar dentro de uma revista e da apresentação desta.

Vê-se que os problemas ligados ao processo de dar forma à mensagem para adaptá-la ao canal e a seus objetivos visados são numerosos e necessitam de técnicas variadas (o estilo, a retórica, a transformação em páginas, em imagens, em cenas, em ondas etc.)

O emissor vive no seu mundo de objetos². O valor de sua mensagem (desconhecida) está nas referências feitas aos objetos contidos na mensagem, em função do que o emissor pensa ser o mundo do receptor. É importante que o canal ofereça um mínimo possível de perturbação e de distorção a esse mecanismo de referência.

Os aspectos fundamentais são evidentes. Eles emprestarão conceitos de outras disciplinas, em particular a semiologia e a lingüística. O conhecimento do contexto da comunicação vai também introduzir impedimentos (sociais, econômicos, jurídicos etc.) que precisarão ser respeitados na concepção da construção do canal. Enfim, a avaliação do bom funcionamento desse canal deve fazer parte das técnicas de medidas cujos fundamentos já são pistas para pesquisas importantes.

Esquema n-1 (n emissores — 1 receptor)

O receptor como ser humano está na impossibilidade de receber n mensagens ao mesmo

tempo. O canal pretendido aqui deverá forçosamente possuir uma memória de estocagem das mensagens. Ele deverá igualmente possuir um dispositivo de seleção permitindo recuperar a informação procurada.

Existe um reservatório de informações e o receptor representa um papel muito mais ativo, tomando a iniciativa de estabelecer a comunicação e participar do funcionamento do canal pela manipulação da seleção. Ele estabelecerá assim a ligação final entre o autor do documento que lhe interessa e si mesmo³.

Os canais desse tipo são: a biblioteca, o centro de documentação, os diferentes tipos de bases de dados (bibliográficas, factuais, numéricas, de conhecimentos etc.).

Fundamentalmente, os problemas se apresentam da mesma maneira como no esquema 1-n precedente. Contudo, eles possuem uma maior precisão, uma vez que os fenômenos de referência podem ser mais fortemente perturbados pelo mecanismo de seleção, como tentaremos mostrar a seguir.

5 - INTERCONEXÕES E ESQUEMAS

Os diferentes casos examinados acima devem ser considerados como esquemas elementares que, combinados entre si, mais ou menos seqüencialmente, podem descrever situações de comunicação complexas.

Assim, o autor de um livro teve, sem dúvida, a necessidade, para escrever, de manter algumas conversações com alguns colegas ou conhecimentos. Ele assistiu a exposições em congressos, que lhe permitiram precisar suas hipóteses. Ele procedeu a algumas pesquisas bibliográficas e leu outras obras.

Seu livro será difundido por um editor e será referenciado em uma base de dados.

Uma tal sucessão de canais implica pontos de ruptura que causam problemas de adaptação onde aparecem ainda os dois níveis técnicos evocados acima:

- as mudanças de suportes físicos;
- a evolução da forma da mensagem.

A passagem de uma língua a outra é um caso típico de adaptação de segundo nível. Mas, a indexação de um documento (mensagem) também o é: é a adaptação da mensagem (documento) ao dispositivo memória-seleção de um canal de tipo n-1.

Todavia, assistimos atualmente a uma certa tendência à homogeneização por intermédio da numerização e da informática. A informação, quer seja veiculada por papel ou por ondas, quer corresponda a uma conversação telefônica, a uma emissão de televisão ou de rádio, encontra-se mais e mais sob uma mesma forma (numérica), que possui uma propriedade entre outras: aquela de se prestar facilmente ao tratamento (é fácil de imaginar as trucagens possíveis sobre uma imagem de televisão numerizada). As possibilidades de ação sobre a forma tornam-se assim cada vez maiores e conduzem a colocar questões fundamentais sobre a forma da mensagem com relação aos objetivos da comunicação e à conservação dos mecanismos de referência.

6-PESQUISAS FUNDAMENTAIS EM CIÊNCIA DA INFORMAÇÃO

Examinar o conjunto de condições de comunicação numa ótica de nível técnico intermediário é uma tarefa imensa e ainda em boa parte por ser feita. É por isso que consideramos apenas os aspectos que correspondem à atividade principal do laboratório e que se referem às bases de dados textuais enquanto canais de comunicação. Iremos examinar sucessivamente a adaptação ao canal na entrada (a integração da mensagem-documento ao sistema de estocagem-seleção) e a adaptação ao receptor na saída.

6.1 - DISPOSITIVO DE SELEÇÃO - LINGUAGEM DE REPRESENTAÇÃO

Integrar uma mensagem-documento consiste em descrever seu conteúdo com a ajuda de uma linguagem de representação, chamada também de linguagem classificatória ou documentária.

Acontece que a concepção de tais linguagens é por demais empírica e fundada sobre critérios essencialmente práticos. Uma sucessão de escolhas arbitrárias preside freqüentemente a construção de uma linguagem documentária. Tais decisões vão intervir⁴:

- no momento de considerar a coleção de termos como fechada;
- no momento de constituir as classes de equivalências detectoras da sinonímia;
- no momento de escolher o termo preferido (o mais curto, o menos ambíguo, ele existe em um tesouro vizinho, existe uma palavra correspondente em uma outra língua etc.);
- na definição de hierarquias, e na escolha de uma delas pretende-se eliminar as poli-hierarquias;

- na constituição de classes de termos religados por uma relação associativa;
- etc.

Parece que o mesmo empirismo reina na construção das classificações⁵.

6.2-ADAPTAÇÃO AO CANAL: A INDEXAÇÃO

A indexação é uma operação intelectual complexa. Um cálculo simples pode facilmente mostrá-lo:

Consideremos um documento de 20 páginas. Ele contém aproximadamente 8 000 ocorrências de palavras, ou seja, em torno de 4 000 palavras diferentes. Se utilizarmos 20 palavras para indexá-lo, teremos uma taxa de condensação fantástica de 0,5%. É, portanto, sobre a escolha desta insignificante proporção de palavras que vai se basear o funcionamento do sistema.

Como o indexador procede? É evidente que ele opera em duas fases:

- Ele lê e compreende o documento. De acordo com seus conhecimentos devido às suas leituras, sua compreensão do documento vai evoluir no tempo;
- Ele emite um discurso, na linguagem de representação, destinado a descrever o conteúdo do documento. Esta descrição tem por objetivo satisfazer um usuário cujas necessidades evoluem igualmente no tempo.

Temos aqui um outro fator (cf. nota 4) de flutuação aleatória da indexação. Suas imperfeições podem ser consideradas como ruído do canal (no sentido usado pelos engenheiros de telecomunicações) e estão na origem das falhas dos sistemas documentários (ruído e silêncio).

6.3-ASPECTOS TEÓRICOS

O aperfeiçoamento das qualidades do canal passa por uma reflexão em profundidade sobre os seguintes aspectos:

- Uma linguagem de representação: qual é sua função exata? Precisa-se mesmo de uma? Se é preciso, como construí-la?;
- Fundamentalmente, o que significa representar o conteúdo de um documento? Qual é o estatuto lingüístico da palavra-chave?;
- Do ponto de vista teórico, em que consiste a operação de interrogar um acervo documentário?

Os problemas são rapidamente recolocados de uma outra forma: descrever o conteúdo do documento é na realidade procurar saber de qual assunto ele trata.

- Será preciso, como preconizam os anglofones e os defensores da inteligência artificial, melhorar as técnicas de representações de conhecimentos? Neste caso, a indexação de um documento torna-se uma projeção sobre um universo construído *a priori*.

- Ou será preciso identificar, graças a analisadores bem construídos, as partes do discurso particularmente representativas dos fenômenos de referências da realidade extralinguística?⁶ Uma confirmação desta última hipótese poderia ser fornecida pela evolução de todos os grandes programas (*softwares*) documentários funcionando sobre um sistema de palavras-chaves, em vista da consideração dos elementos do próprio texto.

A mesma reflexão pode ser feita sobre a operação de pesquisa da informação:

- Será preciso, graças a uma representação dos conhecimentos postos no início, desenvolver inferências permitindo deduzir, a partir de uma questão, os documentos pertinentes?

- Ou será preciso perguntar se a questão "do que fala o documento", tal como foi definido acima, corresponde ao que procura o usuário? Com efeito, as reflexões realizadas no quadro das pesquisas do grupo SYDO-Lyon tendem a mostrar que a introdução de um documento em um fundo documentário faz passar do mundo dos objetos e de um discurso sobre esses objetos (que se situam num quadro de lógica intencional) a um mundo de classe de documentos (regido por uma lógica intencional⁷) Quanto à pesquisa da informação, ela procederia mais pela definição de propriedades que deveriam possuir os documentos pertinentes, e se situaria assim em uma lógica intencional. Em outros termos, a palavra isolada (digamos a palavra-chave) que traduz o enunciado de uma propriedade e não uma referência a um objeto seria um mal instrumento à entrada do canal e bom à saída.

As considerações precedentes mostram bem os problemas fundamentais ressaltados por um exame em profundidade do funcionamento do canal em questão. Pensamos que os trabalhos deste gênero sobre os outros canais são

lucrativos e definem todo um campo de pesquisa constituindo as ciências da informação.

Em particular, apenas evocamos documentos escritos. Conhecemos a importância da imagem e adivinhamos que o mesmo trabalho deveria ser feito sobre este tipo de mensagem. As pesquisas sobre a semiologia da imagem dentro de um contexto de comunicação está ainda dando seus primeiros passos.

Não esqueçamos que a validade das hipóteses sobre o funcionamento do canal deverão ser verificadas e, portanto, que instrumentos de medida apropriados deverão ser criados. Os aspectos fundamentais ligados às medidas e aos instrumentos correspondentes abrem igualmente pistas a importantes pesquisas.

7 - CONCLUSÃO

Vimos que o desenvolvimento de novas tecnologias da comunicação participa da uniformização do suporte elementar e permite um enriquecimento do tratamento. O primeiro nível técnico, tal como o definimos acima, oferece, de mais a mais, a possibilidade de agir sobre a forma da mensagem. Todo o problema está, assim, em definir as transformações necessárias para uma maior eficiência do canal em um dado contexto. Em outros termos, como melhor respeitar as intenções do emissor, que tem um objetivo no momento de produzir sua mensagem, para que sejam percebidas pelo receptor? Como identificar numa mensagem os sinais portadores de referência à realidade extralinguística e veiculá-los o melhor possível? Como criar e organizar esses sinais em função do objetivo da comunicação? Eis aqui, parece-nos, um campo vasto de pesquisa que corresponde a uma seção importante do domínio das ciências da informação.

Para voltar à comunicação em geral, vimos que ela não podia ser concebida sem se levar em conta a dimensão sócio-econômica, que por sua vez introduzira impedimentos (por exemplo, jurídicos, econômicos etc.) que devem ser tratados pelo criador e pelo administrador do canal

¹ Moles, A. *Théorie structurale de la communication*. Paris, Masson, 1986.

² Que podem ser objetos de pensamento.

³ Poderia-se dizer também que o leitor que vem a uma livraria para comprar um livro não tem uma atitude totalmente passiva.

⁴ É relativamente fácil de se ver que nenhuma análise em profundidade da descrição do conteúdo de um documento está na origem desses critérios. De natureza diferente,

e aplicados de acordo com os casos, eles introduzem flutuações aleatórias na indexação que são perturbações dos mecanismos de referência.

⁵ Maniez, J. *Les langages documentaires et classificatoires*. Paris, Les Editions do Organisation, 1987.

⁶ Essas partes do discurso são, assim, as mais representativas do que fala o documento e as mais portadoras de referência.

⁷ Russel, B. *Introduction à la philosophie mathématiques*, Traduzido por Moreau, G. Paris, Payot, 1970.

de comunicação. Vê-se que considerações que, à primeira vista, podiam aparecer como essencialmente técnicas acabam na realidade por serem tratadas sob a forma de pesquisas fundamentais que fazem das ciências da informação um campo científico ainda inteiramente a ser explorado.

INFORMATION SCIENCE: FORM SCIENCE

ABSTRACT

What is the scientific field of information science? In this paper is developed the idea that

information science is concerned by adaptation of the form of the message to a given channel and a fixed goal of communication. Such an aspect, which seems very technical, actually concerns fundamental research. For instance, in the field of written documents, it refers to the problem of reference function to the extra-linguistic reality. Consideration of the same kind can be made about other type of message (fixed image, moving image, radio production, etc.) Consequently, information science appears as a very vast and partly unexplored scientific field.