

**Avanços tecnológicos**

Abrir o espaço semântico em prol da inteligência coletiva

*Pierre Lévy*

Membro da Royal Society of Canada, Canada Research Chair in Collective Intelligence at the University of Ottawa, Ottawa, Canadá
plevy@uottawa.ca

Palavras-chave

Inteligência coletiva, gerenciamento do conhecimento, interoperabilidade semântica, www, cálculo semântico

Resumo e Introdução

Como a memória humana gravada é cada vez mais digitalizada e inserida online, a necessidade de um sistema de coordenadas semânticas independente das línguas naturais e ontologias está aumentando. Um futuro sistema de endereçamento semântico universal, capaz de indexar todos os documentos digitais, deveria atender a três exigências básicas. Primeiro, cada conceito distinto deve ter um único endereço. Segundo, o sistema de coordenadas semânticas deve ser aberto a qualquer conceito e relações entre conceitos (ontologias) independentemente do ambiente cultural em que esses conceitos são criados e transformados, sem privilégios e exclusões. Terceiro, ele deve permitir um grupo de operações matematicamente definidas (possíveis de ser automatizadas) nos endereços semânticos, ou seja: rotações, simetrias e translações em um “espaço semântico”; compressão e descompressão semântica; operações da teoria dos conjuntos como união, interseção e diferenças simétricas; ranking dos critérios semânticos; reconhecimento do padrão semântico; medidas das distâncias semânticas; inferências lógicas etc.

Desenvolvida por uma rede de pesquisa internacional conduzida pela Canada Research Chair in

Collective Intelligence [Cadeira de Pesquisa do Canadá sobre Inteligência Coletiva] da Universidade de Ottawa, a Information Economy MetaLanguage (IEML) [Metalinguagem da Economia da Informação] permite a construção de um sistema de coordenadas semânticas que satisfazem essas três exigências. Site que inclui o Dicionário da IEML desde maio de 2006: <www.ieml.org>. No Brasil, BIREME (www.bireme.br) é membro da iniciativa da IEML.

Interoperabilidade semântica**O problema**

O universo de comunicação aberto, para nós, pela interconexão de dados digitais e manipuladores automáticos de símbolos – em outras palavras, ciberespaço – de agora em diante constitui a memória virtual da inteligência coletiva humana. Todavia, obstáculos importantes impedem que a memória digital funcione completamente em prol de um gerenciamento ótimo dos conhecimentos.

Os obstáculos são:

- a multiplicidade de línguas naturais;
- a incompatibilidade mútua e a adaptação deficiente dos diversos sistemas de indexação e catalogação herdados da era impressa (que não foram concebidos para usar a interconexão e a potência de cálculo do ciberespaço);
- a multiplicidade de ontologias, taxonomias, tesouros, terminologias e classificações;
- as dificuldades encontradas pela engenharia da informação quando tenta levar em conta o *significado* de documentos por meio de métodos genéricos.

Essa série de obstáculos ao desenvolvimento da inteligência coletiva com base digital é conhecida como “problema da interoperabilidade semântica”.

Solução proposta

A metalinguagem da economia da informação (IEML) foi especialmente concebida para resolver este problema. Trata-se de um sistema de digitalização semântica independentemente do formato dos documentos,

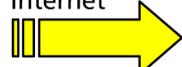
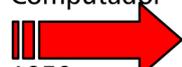
dos sistemas catalogados, das ontologias e das línguas naturais que possibilita identificar, estabelecer relação e manipular automaticamente os *conceitos*.

A IEML permite uma *computação semântica* uniforme quaisquer que sejam os assuntos tratados pelos fluxos e pelos estoques de informações. Ao fazer isso, a metalinguagem abre caminho para um programa de pesquisa técnico-científico que associa as diversas áreas de conhecimento e a informática: gerenciamento do conhecimento computacional. Usada como um dispositivo de endereçamento da memória digital, a IEML possibilita a exploração inteligente e intensiva dos dados, usando métodos *genéricos* (Quadro 1).

As camadas de endereçamento da memória digital

Para compreender a necessidade de uma nova camada de endereçamento da memória no ciberespaço, temos de analisar como se organizam as camadas precedentes.

Quadro 1 – Arquitetura da Memória Virtual

<p>Noosfera 2015</p> 	<p><u>Interconexão entre significados.</u> <IEML> endereços dos conceitos: semiographs </IEML>. <i>Cálculo semântico, exploração multimídia.</i> Memória semântica virtual compartilhada. Inteligência coletiva mundial. Aumento de interpretação</p>
<p>Web 1995</p> 	<p><u>Interconexão entre documentos.</u> URL = http:// endereços de páginas. <i>Ferramentas de pesquisa, navegadores.</i> Esfera pública mundial hipertextual multimídia.</p>
<p>Internet 1980</p> 	<p><u>Interconexão entre computadores.</u> Protocolo de internet = endereços dos servidores de informações. <i>Roteadores.</i> Computadores pessoais. Comunidades virtuais. Convergência de mídias digitalizadas.</p>
<p>Computador 1950</p> 	<p><u>Interconexão entre transistores.</u> Memória de computador = endereços dos bits. <i>Sistemas operacionais. Softwares de aplicativos.</i> Aumento dos processamentos lógicos e aritméticos.</p>

Primeira camada (endereçamento dos bits)

No nível dos computadores que compõem os pontos de ligação no ciberespaço, o sistema local de endereçamento de *bits* de informação é gerenciado de uma maneira descentralizada por vários sistemas operacionais (tais como Unix ou Windows) e utilizado por softwares de aplicativos. O desenvolvimento da informática, na década de 1950, criou condições técnicas para um aumento extraordinário dos processamentos aritméticos e lógicos das informações.

Segunda camada (endereçamento dos servidores)

No nível da rede das redes, cada *servidor* tem um endereço atribuído de acordo com o protocolo universal da internet. Os endereços IP (Internet Protocol) são usados pelo sistema de roteamento – ou comutação – de informações que faz a internet funcionar. O desenvolvimento da internet na década de 1980 corresponde ao advento da informática pessoal, o crescimento de comunidades virtuais e o começo da convergência da mídia e das telecomunicações no universo digital.

Terceira camada (endereçamento das páginas)

No nível da *World Wide Web*, as *páginas* de documentos, por sua vez, ganharam um endereço de acordo com o sistema universal de URLs (*Uniform Resource Locator*) [Localizador Uniforme de Recursos] e os *links* entre documentos são tratados de acordo com a norma HTTP (*HyperText Transfer Protocol*) [Protocolo de Transporte de HiperTexto]. Os endereços da *Web* e os *links* hipertextuais são usados por ferramentas de pesquisa e surfistas da *Web*. A popularização da *Web*, de 1995 em diante, ajudou a dar origem a uma esfera pública mundial multimídia.

Quarta camada (endereçamento dos conceitos)

A Noosfera se apresenta de uma camada adicional da memória digital, com base em um sistema de endereçamento universal dos *conceitos*: IEML. Enquanto sistema de coordenadas da Noosfera, a IEML possibilita o gerenciamento automático das relações entre os conteúdos significativos dos documentos independentemente das línguas e terminologias em que esses documentos foram escritos, catalogados e indexados. O gerenciamento do conhecimento computacional se dedica à manipulação automática dos números semânticos que endereçam os dados dos documentos. Ao fazer isso, ele aumenta a capacidade humana de *interpretação* da memória virtual. Novos dispositivos de *exploração multimídia* do universo dinâmico dos conceitos poderão se apoiar no cálculo semântico.

IEML e a *web* semântica

Ferramentas da *web* semântica

Algumas pessoas podem questionar a necessidade de construir uma nova camada de endereçamento semântico para dados, uma vez que já temos normas e ferramentas da *web* semântica sistematizadas por Tim Berners-Lee. Todavia, a *web* “semântica”, ao contrário do que sugere seu nome, propõe fundamentalmente normas de codificação *lógica* das informações.

As principais ferramentas simbólicas da *web* semântica são:

- XML (*eXtensible Mark-up Language*) [Linguagem de Marcação Extensível], derivada da SGML (*Standard Generalized Mark-up Language*) [Linguagem Padronizada de Marcação Genérica], de Charles Goldfarb; a XML é usada para descrever, de maneira universal, a *estrutura dos bancos de dados*;

- RDF (*Resource Description Framework*) [Estrutura de Descrição de Recursos], que possibilita *catalogar dados* da *web*, junto com a linguagem Sparkl, que pode ser usada para expressar perguntas nos recursos catalogados pela RDF;

- OWL (*Ontology Web Language*) [Linguagem da *Web* para Ontologia], que possibilita descrever ontologias, ou seja, *estruturas conceituais* de várias áreas de conhecimento que podem servir de base para inferências automáticas.

Embora a principal função desses descritores e marcadores seja estimular a pesquisa automatizada de

dados e a execução das operações por robôs de *softwares*, o problema da interoperabilidade não foi resolvido pela *web* semântica, pelo menos na forma de métodos genéricos e ótimos, por no mínimo duas razões: a notação dos conceitos em língua natural é arbitrária e as diversas ontologias são incompatíveis.

Notação alfabética arbitrária dos conceitos em línguas naturais

Mesmo que a XML, a RDF e a OWL formalizem as *relações* entre conceitos na linguagem universal e neutra da lógica, os *próprios conceitos* são anotados por meio de palavras ou abreviações em diferentes línguas naturais. E isto coloca um problema porque (a) existem milhares de línguas naturais diferentes; (b) em cada língua, as palavras podem ter vários significados; (c) o mesmo significado pode ser expresso por várias palavras; sem falar (d) das mudanças de significado devido a variações do contexto e dos pontos de vista.

O sistema de notação dos números, pela posição (seja na base 10, na 2 ou em outra), possibilita uma interpretação universal e inequívoca do significado de cada numeral, e do lugar ocupado por cada numeral na seqüência escrita do número. Assim, o conceito que corresponde à seqüência dos numerais (o número) pode ser automaticamente deduzido dessa seqüência. Ao contrário, a notação alfabética de palavras em línguas naturais leva a códigos arbitrários – encadeamento de caracteres – que podem ser sempre comparados ou ligados a outro encadeamento de caracteres, mas sem que os caracteres ou sua respectiva disposição possam ser interpretados *per se*. Nesse caso, os símbolos básicos representam sons e não elementos de significado.

Em suma, para os autômatos manipuladores de símbolos, os números anotados na ideografia arábica são imediatamente compreensíveis, enquanto as línguas naturais anotadas em letras alfabéticas são semanticamente opacas. Mesmo que as ligações entre as etiquetas lógicas em XML, RDF e OWL sejam calculáveis, os encadeamentos de caracteres que marcam as etiquetas permanecem *códigos arbitrários* do ponto de vista da computabilidade semântica.

A multiplicidade de hierarquias ontológicas

O segundo motivo pelo qual a *Web* semântica não pode resolver sozinha o problema da interoperabilidade semântica é que as ontologias são mutuamente incompatíveis. Geralmente, elas são estruturadas por *hierarquias* de conceitos, e de relações entre conceitos, que permitem que níveis inferiores herdem automaticamente propriedades dos níveis superiores. Essas hierarquias são contextuais, ou seja, estão ligadas a campos da prática ou a escolhas filosóficas e culturais. Obviamente, é possível usar ontologias superiores capazes de organizar um grande número de ontologias locais, como a Cyc, de Douglas Lenat, ou a SUMO¹ do IEEE², e associar a cada conceito de uma ontologia superior sua tradução em um grande número de línguas naturais. No entanto, dificilmente isso resolve o problema da interoperabilidade, pois existem *várias*

ontologias superiores e cada uma delas envolve necessariamente escolhas filosóficas e práticas *particulares*.

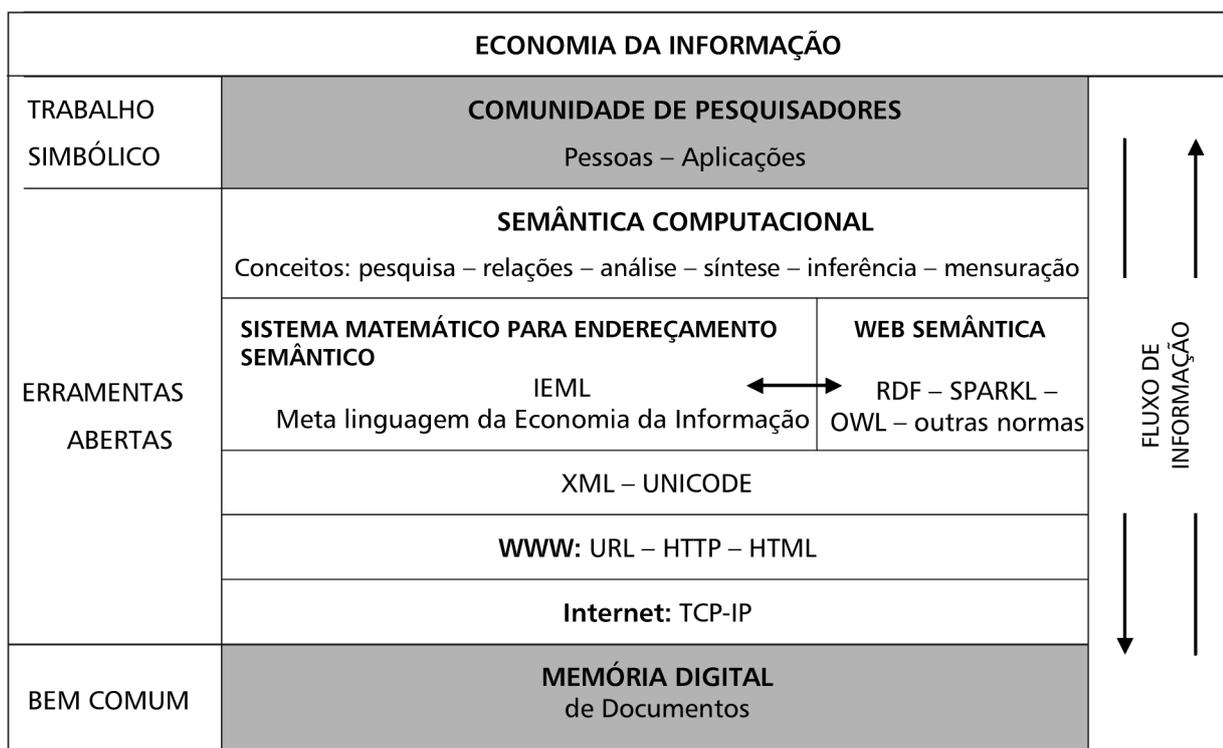
Um sistema universal de endereçamento dos conceitos deve necessariamente: (1) ser independente das ontologias; (2) permitir a expressão de tantas ontologias distintas quanto forem desejadas; e (3) admitir a medida de proximidade entre ontologias, sem privilegiar *a priori* qualquer ponto de vista ontológico.

Complementariedade da IEML e a web semântica

Não se trata, aqui, de colocar em questão a utilidade da XML, RDF ou das ontologias em OWL; mas simplesmente reconhecer que elas não fornecem um sistema matemático universal de endereçamento dos

conceitos. As ferramentas e normas da web semântica são efetivamente necessárias para a implementação técnica de um endereçamento matemático de conceitos, todavia *não são suficientes*. A semântica computacional baseada na IEML possibilitará a valorização das ferramentas e dos métodos produzidos pela *web* semântica, com a qual ela entra em uma relação de complementariedade e valorização recíproca, mais do que de rivalidade. É possível ler e escrever números IEML (semanticamente calculáveis) nas etiquetas lógicas da web semântica, sendo a tradução, em línguas naturais, dos conceitos correspondentes fornecida por um dicionário IEML multilíngüe. Assim, a *web* semântica pode ser considerada um dispositivo lógico intermediário entre a web e a Noosfera. O Quadro 2 mostra a arquitetura técnica geral da economia da informação proposta pela iniciativa da IEML.

Quadro 2 – Economia da informação



Estrutura da IEML

Generalidades

A IEML pode ser considerada uma espécie de ábaco semântico que pode ser manipulado por computadores. Todas as cifras dessa metalinguagem podem ser reconhecidas por uma máquina de estados finitos³. As cifras da IEML são construídas de uma maneira uniforme por meio da geração de fluxos de informações entre um pequeno número de elementos primitivos de acordo com uma hierarquia articulada de níveis estruturais. Uma cifra semântica é sempre construída como um fluxo de informações entre duas ou três cifras semânticas, de ní-

veis de articulação inferiores, que desempenham o papel de fonte, de destino e (eventualmente) de tradutor ou mediador. Uma regra de composição estabelece que o tradutor será vazio se o destino for vazio e que o destino será vazio se a fonte for vazia.

Elementos

A IEML é baseada em cinco elementos básicos, que descrevem os principais componentes do significado:

Elementos verbais (O)

Virtual U e Real A são os dois elementos verbais O, ligados aos processos: (O = U, A).

O *virtual* abrange o universo de possibilidades, coisas a acontecer, potenciais, capacidades, problemas, classes e tipos gerais universais que muito freqüentemente são “intangíveis”. O elemento virtual é caracterizado por uma ausência de coordenadas espaço-temporais.

O *real* ocorre no tempo e no espaço. É constituído por indivíduos singulares, eventos originais, formas surgidas, soluções de problemas, exemplos típicos de fenômenos universais e dados perceptíveis.

A dialética da ação (O) organiza uma troca de informações, uma circulação de diferenças entre o virtual e o real: cada atualização transforma o virtual e cada transformação do virtual gera uma nova realidade.

A dialética do virtual (U) e do real (A) é encontrada em um grande número de tradições filosóficas e culturais: o céu e a terra das primeiras filosofias, a transcendência e a imanência das teologias, o yin e o yang do taoísmo, o inteligível e sensível do platonismo, o numenal e fenomenal de Kant, o vácuo e os fenômenos do budismo maaiana etc.

Na maioria de suas definições abstratas, o virtual define um domínio de variação e o real um operador: a combinação dos dois papéis cria uma função.

Elementos nominais

Signo S (significante), ser B (significado para um interpretante) e coisa T (referente) são os três elementos *nominais* M, ligados às representações (M = S, B, T).

O *signo* corresponde ao significante em lingüística. É um instrumento simbólico cuja operação principal é apontar para os referentes do discurso humano. Signos são os sons da palavra, os caracteres da escrita, gestos e sinais, imagens e sinais de todos os tipos, geralmente símbolos que podem ser interpretados. “O dedo aponta para a lua. O idiota olha para o dedo”, como diz um provérbio zen. Neste exemplo, o dedo representa o significante (em outras palavras, o *signo*) S, enquanto a lua é o referente (em outras palavras, a *coisa*) T.

Agora, com exceção dos nomes próprios que designam realidades singulares, é impossível ligar um significante a uma referência singular da fala, sem primeiro passar por um conceito intermediário associado ao signo:

o significado. Por sua vez, o significado somente pode significar *para um interpretante*. Esse significado, que é indissociável de seu interpretante subjetivo, é denominado *ser* na IEML. O *ser* completa o movimento cognitivo que passa do dedo (o *signo*) para a lua (a *coisa*) e dá um valor contextual para essa relação signo-referente.

Os elementos nominais da IEML (*signo* S, *ser* B, e *coisa* T) são os três fatores de representação distintos e interdependentes. Mas atenção: eles são diferenciados por sua função, e não por sua natureza intrínseca. Dependendo das várias perspectivas cognitivas, uma pessoa, por exemplo, pode desempenhar o papel de *signo* (o significante do discurso) ou de *ser* (o interpretante do discurso) ou de *coisa* (o objeto do discurso).

Na dialética semântica, o signo, o ser e a coisa foram denominados *vox*, *conceptus* e *res* na universidade medieval. Na filosofia de C. S. Peirce, estes foram traduzidos como signo (ou *representamen*), interpretante e objeto. Suas variações na lingüística moderna são o significante, significado e referente. Essa dialética semântica é encontrada na lógica (proposições, julgamentos, circunstâncias), na economia (preço, propriedade, utilidade) e na teologia (ensinamentos, comunidade, realidade final). Essa dialética ternária pode ser detectada no *trivium* das artes liberais na Antiguidade e na Idade Média ocidental: a *gramática* desenvolve o domínio da linguagem (a manipulação de signos), a *dialética* oferece uma introdução ao diálogo racional (entre seres), a retórica se ocupa da construção prática do discurso visando à sua memorização e aos efeitos reais (sobre as coisas).

Assim, as primitivas da IEML – uma linguagem do endereçamento de dados numéricos de acordo com o seu significado – são, como era de se esperar, as próprias estruturas do significado. Essas estruturas foram descritas por antigas e inúmeras tradições pertencentes a várias culturas e disciplinas. Simplesmente, contentei-me em recolhê-las e relacioná-las.

De eventos a frases

A partir desses cinco elementos, a IEML desenvolve quatro níveis de combinação e articulação das cifras semânticas (Quadro 3):

Quadro 3 - Alfabeto IEML

OO energias	OM ações	MO mutações	MM conceitos
U → U w o refletir	U → S y saber	S → U j significante	S → S s pensamento
U → A w a agir	U → B o querer	B → A g documentário	S → B b linguagem
A → U w u perceber	U → T e poder	B → U h significado	S → T t memória
A → A w e reconstituir	A → S u expressar	B → A c pessoal	B → S k sociedade
	A → B a comprometer-se	T → U p referente	B → B m afeto
	A → T i fazer	T → A x material	B → T n mundo
			T → S d verdade
			T → B f vida
			T T l espaço

- 25 (5^2) eventos, ou “letras semânticas”, que são fluxos de informações entre dois elementos (os eventos são representados por 25 letras minúsculas em negrito no quadro acima, em que as vogais são verbos e as consoantes são substantivos);

- 625 (25^2) relações, que são combinações de duas letras ou “sílabas semânticas” da metalinguagem;

- 240 milhões de idéias ($625^2 + 625^3$), combinações de 4 ou 6 letras, que representam as “palavras” da metalinguagem;

- uma quantidade astronômica (10^{23}) de “frases” que combinam uma, duas ou três idéias.

Em suma, a gramática da IEML tem cinco níveis de articulação: elementos, eventos, relações, idéias e frases.

O espaço semântico

Dimensões e perspectivas

Considera-se que o espaço semântico endereça uma quantidade praticamente infinita de diferentes grafos de frases de IEML. Matematicamente, um grafo pode ser definido por uma série de triplos. Cada triplo é composto por 1) um vértice inicial; 2) um vértice de chegada; 3) uma ligação entre os dois vértices. A frase da IEML que etiqueta o vértice inicial é denominada frase fonte (So), a frase da IEML que etiqueta o ponto de chegada é denominada frase destino (De) e a frase da IEML que etiqueta a ligação é denominada frase tradutora (Tr). Os triplos que compõem os grafos das frases da IEML podem ser representados como “pontos” abstratos de um espaço 3D [tridimensional] do qual os três eixos são So, De, Tr.

1ª dimensão: fonte; 2ª dimensão: destino; 3ª dimensão: tradutor. Em cada dimensão, as variáveis são as 10^{23} frases da IEML.

Assim, o espaço semântico da IEML é uma matriz cúbica abstrata que contém 10^{69} unidades básicas, ou pixels semânticos que são *triplos* das frases da IEML.

O espaço *semântico* 3D pode ser projetado em muitos espaços 3D *geométricos* denominados perspectivas semânticas. Há tantas perspectivas semânticas quantas ordens estritas existirem entre frases nos três eixos do espaço semântico. Qualquer ordem possível de frases ao longo dos três eixos produz uma diferente projeção geométrica 3D do espaço semântico. Uma perspectiva semântica não se baseia em um ponto do espaço 3D mas em um espaço 3D completo de uma matriz de espaços 3D possíveis (o espaço semântico).

Endereçamento no espaço semântico

Como vimos anteriormente, o “pixel semântico” ou unidade básica do espaço semântico é um triplo de frases da IEML (So, De e Tr). Esta unidade é denominada um número semântico. Existem 10^{69} triplos de frases ou números semânticos (na verdade, um pouco menos do que isso porque é necessário um destino não-vazio para se conseguir um tradutor não-vazio e uma fonte não-vazia

para se obter um destino não-vazio). Projetado em uma perspectiva semântica, o número semântico torna-se um ponto geométrico.

Um grafo de frase da IEML é uma série de numerais semânticos e define uma subsérie do espaço semântico. É denominado número semântico. Existe uma quantidade astronômica de números semânticos possíveis. Mesmo que teoricamente finito, ele é praticamente infinito.

Projetado em uma perspectiva semântica, o número torna-se uma série de pontos geométricos, uma “figura”. Os números semânticos (coordenadas do espaço semântico) são comuns a todas as perspectivas semânticas: a única diferença é sua projeção 3D em uma figura, que é ligada a uma perspectiva semântica particular.

Dados semânticos

Os dados semânticos representam conceitos aos quais se atribui valor e referência. Um dado semântico é composto por três partes: 1) o conceito formal ou endereço semântico; 2) os valores do conceito; 3) as referências do conceito.

1) A coordenada espacial única de um conceito é dada por um número semântico, ou seja, por uma série de numerais semânticos ou, em outras palavras, por uma subsérie do espaço semântico.

2) Os valores do conceito correspondem à ordenação (ou ranking) de números associados a um número semântico e às quantidades – ou números cardinais – associados a um número semântico. Números ordinais dependem de funções de ranking explícitas (se não automáticas) e os números cardinais dependem de funções de mensuração explícitas (se não automáticas). Vários valores podem ser associados ao mesmo conceito formal, de acordo com as várias funções de atribuição de valor.

3) As referências são links para endereços físicos de documentos (URLs, por exemplo). Vários endereços físicos de documentos podem ser associados ao mesmo número semântico; por exemplo, documentos com conteúdo semântico equivalente, mas em diferentes línguas naturais. Cada endereço físico de documento depende de uma função de indexação explícita (se não automática). O mesmo endereço físico pode ser associado a diferentes coordenadas semânticas de acordo, por exemplo, com diferentes funções de indexação. Finalmente, um endereço físico pode conter um endereço semântico (auto-referência do espaço semântico).

Cálculo dos dados semânticos

Dados semânticos são compostos por três partes diferentes: endereço, valores e referências de um conceito. Duas delas, o endereço e os valores, podem sempre admitir manipulações automáticas porque são compostas por números. O endereço do conceito – ou conceito formal – é um número semântico que pode ser manipulado por uma máquina de estados finitos. Os valores são números ordinais (que dependem de funções de ranking) e números cardinais (que dependem de funções de mensuração). Portanto, é sempre possível definir funções calculáveis.

Obs.: a parte da referência dos dados semânticos depende das funções de indexação que nem sempre são calculáveis como, por exemplo, as convenções que resultam do acordo de um grupo de intérpretes humanos que são registradas nos dicionários de línguas naturais da IEML. No entanto, é possível programar autômatos de indexação a partir dos dicionários.

Entre as várias funções dos dados semânticos que podem ser automaticamente calculadas, citaremos:

- rotações, translações e simetrias de conceitos formais no espaço semântico;
- operações da teoria dos conjuntos (união, interseção, diferenças simétricas) dos conceitos formais;
- compressões e descompressões (síntese e análise) de conceitos formais a partir de classificações;
- ranking automático de conceitos formais de acordo com a semântica ou com critérios externos;
- funções verdadeiras (valor 0 ou 1);
- reconhecimento dos padrões/normas semânticos(as).

A composição de funções define *autômatos semânticos* que refletem os interesses, as interpretações e as operações cognitivas de uma comunidade de intérpretes e aumenta sua capacidade de gerenciamento do conhecimento.

Significado dos conceitos formais

Atribuição de descritores em línguas naturais aos vértices semânticos da IEML

As línguas naturais são múltiplas, ambíguas e sujeitas a mudanças. Portanto, é impossível *deduzir automaticamente* a interpretação dos vértices semânticos da IEML (principalmente idéias e frases) em línguas naturais. Essa interpretação pode ser apenas *convencional*. Em compensação, uma vez dada a interpretação das idéias da IEML (as “palavras” da metalinguagem), a interpretação dos números semânticos (os “textos” da metalinguagem: grafos de frases) pode ser gerada automaticamente.

Uma vez que o objetivo da metalinguagem é calcular automaticamente relações semânticas, a atribuição de descritores em línguas naturais às idéias e frases da IEML não pode ser arbitrária: na medida do possível, ela deve se conformar aos três principais critérios abaixo:

Critério de simetria. As *simetrias semânticas* da metalinguagem devem se refletir nas *simetrias semânticas* reveladas pelos descritores em línguas naturais.

Critério de economia. A atribuição dos descritores deve possibilitar gerar, por composição, um *máximo* de conceitos por meio de um *mínimo* de símbolos da IEML.

Critério de composição. A *interpretação de uma combinação de símbolos da IEML* por descritor em língua natural deve corresponder tanto quanto possível (nem sempre é possível) à *combinação das interpretações desses símbolos*.

Com o objetivo de iniciar o processo de interpretação dos conceitos formais da IEML, o autor traduziu em

línguas naturais (francês e inglês) as 625 relações e mais de mil idéias abrangendo a maioria dos objetos e disciplinas das ciências humanas (em que é especializado). Agora, é necessário que a continuação do processo de interpretação seja um empreendimento coletivo aberto, do qual são convidados a participar: (a) os gestores de ontologias, terminologias, tesouros e classificações; (b) os especialistas das áreas de conhecimento que quiserem formalizar seus conceitos em IEML; e (c) os tradutores que estão desenvolvendo o dicionário IEML. A principal ferramenta desse processo de interpretação conjunto é um dicionário wiki denominado “wikimetal” (wiki da metalinguagem) que pode ser encontrada no site <www.ieml.org> desde abril de 2007.

Polissemia

Na IEML, um conceito formal (um número semântico) é unívoco: cada endereço do espaço semântico é único, distinto e sem ambigüidade. Apesar disso, a IEML não foi inventada para eliminar mas, ao contrário, para *aumentar* as possibilidades contextuais de interpretação.

No espaço semântico, a multiplicação de interpretações (ou polissemia) não se baseia na ambigüidade de conceitos, mas na imensa variedade de *operações* (transformação, ranking, mensuração e indexação) que podem ser desempenhadas nos conceitos. *Portanto, a multiplicidade de contextos geradores de sentido* é moldada pela multiplicidade de autômatos semânticos capazes de compor suas operações em um encadeamento semântico. Existem tantos autômatos semânticos possíveis (contextos geradores de sentido) quantas comunidades de intérpretes possíveis.

Conclusão: a interdependência dos três problemas tratados pela semântica computacional

Como vimos, a semântica computacional baseada na IEML propõe tratar o problema da interoperabilidade no ciberespaço. Para concluir este artigo, gostaria de enfatizar a interdependência entre as soluções para três problemas: (1) o problema da interoperabilidade semântica; (2) o problema do apoio para a tomada de decisão do gerenciamento do conhecimento nas organizações; (3) o problema do estudo científico dos processos da inteligência coletiva humana.

IEML e interoperabilidade semântica

A solução para o problema da interoperabilidade supõe o uso da metalinguagem, que é: (a) capaz de dar endereços únicos a conceitos distintos; (b) manipulável por computadores; e (c) capaz de traduzir as várias línguas naturais, ontologias e sistemas de classificação que, hoje, fragmentam a *indexação de documentos na web*. A necessidade dessa metalinguagem está começando a ser reconhecida pela comunidade técnico-científica que gravita em torno da web semântica. Uma das repercussões mais óbvias da adoção desse sistema de coordenadas semânticas universais seria abrir o caminho para *ferramen-*

tas de pesquisas semânticas personalizadas, trabalhando com conceitos em vez de encadeamentos de caracteres.

As ferramentas de pesquisas semânticas são caracterizadas pela capacidade de:

- a) produzir ranking automático de resultados a partir de critérios semânticos;
- b) calcular distâncias geométricas entre padrões conceituais de acordo com sensores semânticos personalizados;
- c) gerar automaticamente síntese, análise e inferências lógicas através de ontologias.

IEML e gerenciamento do conhecimento

Não só a IEML é capaz de traduzir mutuamente as várias línguas naturais e ontologias, como também constitui uma ferramenta para representar e simular os vários *ecossistemas de conceitos* mantidos pelas coletividades humanas (negócios, escolas, universidades, disciplinas, entidades territoriais, associações e comunidades virtuais de todos os tipos). Uma vez que os ecossistemas de conceitos são representados em uma metalinguagem padrão, os dados semânticos podem ser acumulados e comparados, e é possível desenvolver um gerenciamento do conhecimento científico.

A IEML é concebida para *auxiliar a tomada de decisão* no gerenciamento do conhecimento, baseada em uma explicitação do objetivo da comunidade usuária e em uma representação o mais matizada possível da dinâmica conceitual existente (e não como uma função de métodos ou teorias *a priori*). Assim, a semântica computacional é chamada a orquestrar o desenvolvimento inovador do conhecimento em tempo real e a coordenação prática da competência de grupos e coletivos de todas as espécies e escalas.

IEML e a observação científica da inteligência coletiva

Um vez que o problema da interoperabilidade semântica for resolvido, por meio de uma metalinguagem capaz de representar e simular ecossistemas de conceitos, será possível observar cientificamente os processos da inteligência coletiva humana. Na verdade, o volume de memória cultural acumulada assim como uma proporção crescente da comunicação e das transações humanas pairam no universo digital online. Assim, teoricamente é possível usar o ciberespaço como instrumento para observar a inteligência coletiva humana desde a escala de pequenos grupos até a escala mundial. Todavia, para que essa possibilidade se realize de acordo com as expectativas, deveremos primeiro ser capazes de distribuir e situar os fluxos e estoques de informações em um espaço semântico unificado, um espaço que seja capaz de acomodar uma variedade ilimitadamente aberta de conceitos em interação e transformação. Dentro dessa perspectiva, a IEML apresenta-se como um sistema de localização (ou endereçamento científico) de conceitos que possibilita abrir o espaço semântico – como uma *natureza* da mente – para a observação científica. E essa observação inevitavelmente terá repercussões epistemológicas importantes

nas ciências humanas e sociais, assim como nas aplicações práticas em benefício do desenvolvimento humano. Neste sentido, a semântica computacional baseada na IEML pode ser compreendida como uma disciplina auxiliar das ciências humanas.

A inseparabilidade dos três problemas

Em suma:

- a) a idéia de uma linguagem comum para a web está começando a progredir;
- b) a **jovem disciplina de gerenciamento do conhecimento está buscando e descobrindo teorias científicas, métodos e ferramentas;**
- c) **nos últimos 15 anos, a pesquisa – e o discurso teórico – sobre a inteligência coletiva tem sido ampliada.**

A semântica computacional baseada na IEML faz com que essas três correntes de pesquisa compartilhem uma série de equipamentos de cálculos matemáticos, mensuração e endereçamento conceitual. Esses equipamentos podem: (1) resolver o problema da interoperabilidade semântica; (2) oferecer um padrão para a representação de ecossistemas de conceitos e servir de ajuda para a tomada de decisão no gerenciamento do conhecimento; (3) servir de base para a construção de um instrumento de observação científica organizada dos processos de inteligência coletiva.

Nenhum dos três problemas pode ser perfeitamente resolvido a menos que os outros dois o sejam. Qualquer tentativa de solucioná-los *separadamente* pode levar apenas a resultados parciais ou deficientes. A ocasião de um salto na inteligência coletiva iria se perder se a linguagem comum da web (que necessariamente será construída em um prazo relativamente prazo, e sob pressão da necessidade) não desse acesso à observação de um espaço semântico ainda invisível e, ao mesmo tempo, não possibilitasse a organização, com o auxílio do computador, do gerenciamento científico do conhecimento em prol do desenvolvimento humano.

Notas

1. Suggested Upper Merged Ontology. Ontologia superior, criada por Ian Niles, Adam Pease (principal pesquisador) e Chris Menzel (colaborador). Considerada a maior ontologia pública formal existente atualmente, é usada para pesquisa e aplicativos de busca, de lingüística etc. (N.T.)
2. Institute of Electrical and Electronic Engineers (Instituto de Engenheiros Eletricistas e Eletrônicos). (N.T.)
3. Ou autômatos finitos. Para uma definição, ver <<http://pt.wikipedia.org/wiki/Aut%C3%B4mato>> (N.T.)

Referências bibliográficas

Biologia, teoria da evolução

ATLAN, H. *Les Etincelles de hasard, T. 1: Connaissance spermatique*. Paris: Seuil, 1999.

ATLAN, H. **Les Etincelles de hasard, T. 2: Atheisme de l'écriture.** Paris: Seuil, 2003.

ATLAN, H. **Entre le cristal et la fumée.** Paris: Seuil, 1979.

ATLAN, H. **L'Organisation biologique et la théorie de l'information.** Paris: Hermann, 1972.

DAWKINS, R. **The Blind Watchmaker.** NY: Norton, 1986.

KUPIEC, J.; SONIGO, P. **Ni Dieu ni Gène.** Paris: Seuil, 2000.

LANGTON, C.G. (ed.). **Artificial Life**, Santa Fe Institute Studies in the Sciences of Complexity Proceedings, vol.6, Addison Wesley, Redwood, Calif., 1989.

LANGTON, C.G. (Ed.). **Artificial life: An overview.** Cambridge, Mass: MIT Press, 1995.

LOVELOCK, J. **The Ages of Gaïa.** NY: Norton, 1988.

SONIGO, P.; STENGERS, I. **L'Evolution.** Les Ulis (Fr.): EDP Sciences, 2003

TEILHARD DE CHARDIN, P. **Le Phénomène humain.** Paris: Seuil, 1955

Desenvolvimento humano

ANDERSON, M.B. **Development and Social Diversity.** Oxford: Oxfam UK, 1999.

CRUSH, J. (ed.). **Power of Development.** London and New York: Routledge, 1995.

DIEHL, P.F. **The Politics of Global Governance: International Organizations in an Interdependent World.** Boulder, Colorado: Lynne Rienner Publishers, 2001.

McMICHAEL, P. **Development and Social Change: a Global Perspective.** Thousand Oaks, California: Pine Forge Press, 2000.

SEN, A. **Development as Freedom.** New York: Anchor Books, 1999.

UNITED NATIONS DEVELOPMENT REPORT. **Human Development Report 2002.** New York and Oxford: Oxford University Press, 2002.

Economia do conhecimento

AXELROD, R. **The Evolution of Cooperation.** New York: Basic Books, 1984.

CASTELLS, M. **The Information Age, Economy, Society and Culture.** 3 vol. Oxford: Blackwell, 1998.

DIGITAL OPPORTUNITY INITIATIVE (Final Report). **Creating a Development Dynamic,** NY: United Nations, 2001.

HAYEK, F. **Economics and Knowledge**, in *Economica IV* (new ser., 1937), p.33-54.

HAYEK, F. **The Use of Knowledge in Society**, *American Economic Review*, XXXV, n.4; Sep., 1945, p.519-30

KELLY, K. **Out of Control.**, *The New Biology of Machines, Social Systems and the Economic World.* NY: Addison Wesley, 1994.

LAZZARETTO, M. **Puissances de l'invention**, Les empêcheurs de penser en rond-Seuil. Paris, 2002.

MACHLUP, F. **Knowledge, its Creation, Distribution and Economic Significance.** 3 vol, Princeton UP, 1984.

PORAT, M.; RUBIN, M. **The Information Economy**, Government Printing Office, Washington DC, 1977.

STEWART, T.A. **Intellectual Capital: The New Wealth of Organizations.** NY: Doubleday/Nicholas Brealy, 1997.

Gestão do conhecimento

GIBBONS, M. **The New Production of Knowledge: the Dynamics of Science and Research in Contemporary Societies.** London: Sage, 1994.

LEVINE, R.; et al. **The Cluetrain Manifesto, the End of Business as Usual.** Cambridge, Mass.: Perseus Books, 1999.

MOREY, D.; MAYBURY, M.; THURASINGHAM, B. (Ed.). **Knowledge Management, Classic and Contemporary Works.** Cambridge, Mass.: MIT Press, 2000.

BROWN, J.S.; DUGID, P. **The Social Life of Information**, Harvard Business School Press, 2000.

SENGE, P. **The Fifth Discipline: The Art and Practice of the Learning Organisation.** Random House, 1990.

SVENONIUS, E. **The intellectual Foundation of Information Organization.** Cambridge, Mass.: MIT Press, 2000.

WENGER, E. **Communities of Practice: Learning, Meaning, and Identity.** Cambridge: Cambridge University Press, 1998.

Sociedade em rede e capital social

BARABASI, A.L. **Linked, the New Science of Networks.** Cambridge, Mass.: Perseus publishing, 2002.

BUCHANAN, M. **Nexus: Small Worlds and the Groundbreaking Science of Networks.** Norton and Cie, 2002.

GROSS STEIN, J. et al. **Networks of Knowledge. Collaborative Innovation in International Learning.** Toronto: University of Toronto Press, 2001.

LIN, N. **Social Capital: A Theory of Social Structure and Action.** Cambridge, Mass.: Cambridge University Press, 2001.

LUHMANN, N. **Social Systems.** CA: Stanford UP, 1995.

SCOTT, J. **Social Network Analysis**, sec. edition, London: SAGE Publications, 2000.

PUTNAM, R.D. **Bowling Alone: The Collapse and Revival of American Community**. New York: Simon & Schuster, 2000.

WASSERMAN, S.; FAUST, K. **Social Network Analysis**. Cambridge, Mass.: Cambridge University Press, 1994.

WATTS, D.J. **Six degrees, the science of a connected Age**. NY: Norton, 2002.

Ciências cognitivas

BATESON, G. **Steps to an Ecology of Mind**. 2 vol., NY: Chandler, 1972.

BROOKS, R.A.; STEELS, L. (eds.). **The Artificial Life Route to Artificial Intelligence: Building Embodied Situated Agents**. NJ: Lawrence Erlbaum Associates, Inc., Hillsdale, 1995.

EDELMAN, G. **Neural Darwinism**. New York: Basic Books, 1987.

JACKENDOFF, R. **Languages of the Mind**. Cambridge, Mass.: MIT Press, 1992.

JOHNSON-LAIRD, P.N. **Mental Models**. England: Harvard University Press, 1983.

MATURANA, H.R.; VARELA, F.J. **The Tree of Knowledge: The Biological Roots of Human Understanding**. Boston: New Science Library, 1988.

McLELLAND, J.L.; RUMELHART, D.E. and the PDP Research Group, **Parallel Distributed Processing: Explorations in the Micro-Structure of Cognition**, 2 vol. MIT Press, 1986.

MINSKY, M. **The Society of Mind**. NY: Simon and Schuster, 1997.

SOWA, J.F. **Conceptual Structures: Information Processing in Mind and Machine**. MA.: Addison-Wesley, Reading, 1984.

SOWA, J.F. **Knowledge Representation: Logical, Philosophical, and Computational Foundations**. CA: Brooks/Cole Publishing Co., Pacific Grove, 2000.

VARELA, F. **Autonomie et connaissance**. Paris: Seuil, 1989.

VARELA, F.; THOMPSON, E.; ROSH, E. **The Embodied Mind: Cognitive Science and Human Experience**. Cambridge, Mass.: MIT Press, 1991.

Sinais e linguagem

AUSTIN, J.L. **How to Do Things With Words**. Oxford U. P. 1962.

BICKERTON, D. **Language and Human Behavior**. Seattle: University of Washington Press, 1995.

De LIBERA, A. **La Querelle des universaux, De Platon à la fin du Moyen-Age**. Paris: Seuil, 1996.

ECO, U. **Segno**. Milan: ISEDI, 1973.

LANGACKER, R.W. **Foundations of Cognitive Grammar, Vol I**, Stanford University Press, 1987.

PANNACCIO, C. **Les Mots, les concepts et les choses, la sémantique de Guillaume d'Occam et le nominalisme d'aujourd'hui**. Paris-St Laurent (Qc): Bellarmin-Vrin, 1992.

VIGOTSKY, L. **Thought and Language**. Cambridge, Mass.: MIT Press, 1986. (1ère édition russe: 1934)

WITTGENSTEIN, L. **Philosophical Investigations**, Oxford: Blackwell, 1958.

Pioneers of cybernetics and cyberculture

ASHBY, R. **Introduction to Cybernetics**. London: Methuen, 1956.

BARDINI, T. **Bootstrapping. Douglas Engelbart, Co-evolution and the Origins of Personal Computing**. Stanford UP, 2000.

BUSH, V. **As We May Think**; in *The Atlantic Monthly*; Jul., 1945; v.176, n.1; p.101-108 (reproduced in Packer and Jordan, 2001)

ENGELBART, D. **Augmenting Human Intellect**, Technical Report, Stanford Research Institute, 1962.

FOERSTER (von), H. **Observing Systems**. Seaside CA: Intersystems, 1981.

LÉVY, P. **Analyse de contenu des travaux du Biological Computer Laboratory (BCL)**, in *Cahiers du CREA*, 8, Paris, 1986, p.155-191.

LÉVY, P. **L'Oeuvre de Warren McCulloch**, in *Cahiers du CREA*, 7, Paris, 1986, p.211-255.

LÉVY, P., **Wittgenstein et la Cybernétique**, in *Cahiers du CREA*, 7, Paris, 1986, p.257-285.

LICKLIDDER, J.C.R.; TAYLOR, Robert. **The Computer as a Communication Device**, *Science and Technology*, Apr. 1968.

LICKLIDDER, J.C.R. **Man-computer Symbiosis**, *IRE Transactions on Human Factors in Electronics*, v. HFE-1, p.4-11, Mar. 1960.

McCULLOCH, W. **Embodiments of Mind**. Cambridge: MIT Press, 1965.

NELSON, T.H. **Literary Machines**. Mindful Press, 1982.

SHANNON, C.E.; WEAVER, W. **Mathematical Theory of Communication**, U. of Illinois Press, Urbana, 1964.

TOFTS, D.; JONSON, A.; CAVALLARO, A. (Eds), **Prefiguring Cyberculture, an Intellectual History**. Cambridge: MIT Press, 2003.

WIENER, N. **The Human Use of Human Beings: Cybernetics and Society**. NY: Doubleday, 1950.

Inteligência coletiva na cibercultura

AUTHIER, M.; LÉVY, P. (préface de Michel Serres). **Les Arbres de connaissances**. Paris: La Découverte, 1992.

HEYLIGHEN, F. The World-Wide Web as a Super-Brain: From Metaphor to Model, in: **Cybernetics and Systems '96**, R. Trappl (ed.), World Science, Singapore, 1998.

HIMANEN, P. **The Hacker Ethic and the Spirit of the Information Age**. NY: Random House, 2001

JOHNSON, S. **Emergence, The Connected Lives of Ants, Brains, Cities and software**. NY: Scribner, 2001.

LÉVY, P. **L'Intelligence collective. Pour une anthropologie du cyberspace**. Paris: La Découverte, 1994. English translation: LÉVY, P. **Collective Intelligence: Mankind's Emerging World in Cyberspace**. Cambridge, Mass.: Perseus Books, 1997.

LÉVY, P. **Cyberculture**. Paris: Odile Jacob, 1997. English translation: LÉVY, P. **Cyberculture**. Minneapolis: University of Minneapolis Press, 2001.

LÉVY, P. **Cyberdémocratie: Essai de philosophie politique**. Paris: Odile Jacob, 2002.

LÉVY, P. **De la programmation considérée comme un des beaux-arts**. Paris: La Découverte, 1992.

LÉVY, P. **La Machine Univers. Création, cognition et culture informatique**. Paris: La Découverte, 1987.

LÉVY, P. **Les Technologies de l'intelligence. L'avenir de la pensée à l'ère informatique**. Paris: La Découverte, 1990.

LÉVY, P. **Qu'est-ce que le virtuel?** Paris: La Découverte, 1995. English translation: LÉVY, P. **Becoming Virtual. Reality in the Digital Age**. New York, Plenum Trade, 1998.

LÉVY, P. **World Philosophie: le marché, le cyberspace, la conscience**. Paris: Odile Jacob, 2000.

MAYER-KRESS, G.; BARCZYS, C., The Global Brain as an Emergent Structure from the Worldwide Computing Network, and its Implications for Modeling, **The Information Society**, v.11, n.1 (Jan.-Mar. 1995), p.1-27.

RHEINGOLD, H. **Smart Mobs, The next social Revolution**. Cambridge Mass.: Perseus Books, 2002.

ROSNEY (de), J. **L'homme symbiotique**. Paris: Seuil, 1995.

RUSSELL, P. **The Global Brain Awakens: Our Next Evolutionary Leap**, 1996.

WEINBERGER, D. **Small Pieces Loosely Joined, a Unified Theory of the Web**. Cambridge Mass.: Perseus Books, 2002

Web semântica

BERNERS LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities**, *Scientific American*, May 2001.

BERNERS LEE, T. Semantic Web Road map, 1988. Disponível em: <<http://www.w3.org/DesignIssues/Semantic.html>>.

BERNERS LEE, Tim. **Weaving the Web**. San Francisco: Harper, 1999.

FENSEL, D.; et al. (Eds.), **Spinning the Semantic Web, Bringing the World Wide Web to Its Full Potential**. Cambridge: MIT Press, 2003

<http://www2007.org/>

www.semanticweb.org

www.W3.org

www.ieml.org

Visualização e mapeamento

DODGE, M.; KITCHIN, R. **Mapping Cyberspace**. London: Routledge, 2001.

KITCHIN, R.; Freundschuh, S. (eds.). **Cognitive Mapping: Past, Present and Future**. NY: Routledge, 2000.

LÉVY, P. **L'idéographie dynamique. Vers une imagination artificielle?** Paris: La Découverte, 1991.

SHNEIDERMAN, B.; BEDERSON, B.B. (eds.). **The Craft of Information Visualization**. San Francisco, California: Morgan Kaufman Publishers, 2003.

SPENCE, R. **Information Visualization**. Essex, England: ACM Press, 2001.

WARE, C. **Information Visualization: Perception for Design**. San Diego, CA: Academic Press, 2000. 

Sobre o autor

Pierre Lévy

Filósofo que dedicou sua vida profissional ao entendimento das implicações culturais e cognitivas das tecnologias digitais promovendo o seu melhor uso social e ao estudo do fenômeno da inteligência coletiva humana. Escreveu vários livros sobre este assunto que foram traduzidos para mais de 12 idiomas e são utilizados em diversas universidades no mundo inteiro. Atualmente leciona no departamento de comunicação da Universidade de Ottawa (Canadá), onde tem o cargo de Presidente de Pesquisa em Inteligência Coletiva do Canadá (Canada Research Chair in Collective Intelligence). Seu livro "As Tecnologias da Inteligência", publicado no início dos anos 90, previu o advento da Web. Em 1992, fundou na França a primeira empresa de software dedicada ao gerenciamento de conteúdo. Seu livro sobre inteligência coletiva, publicado em 1994, ainda inspira jovens pesquisadores. Ele é o autor de uma linguagem artificial (www.ieml.org) capaz de expressar qualquer conceito de forma computável. IEML poderia se tornar o <<código semântico>> do espaço cibernético, oferecendo à inteligência coletiva humana uma memória compartilhada interoperável. Pierre Lévy é membro do Royal Society of Canada e já recebeu várias premiações e distinções acadêmicas.