



**RECIIS**

Revista Eletrônica de Comunicação  
Informação & Inovação em Saúde

[www.reciis.cict.fiocruz.br]

ISSN 1981-6278

**Artigos originais**

# Bases ontológicas e conceituais para um modelo do conhecimento científico em artigos biomédicos

DOI: 10.3395/reciis.v3i1.240pt



*Carlos Henrique  
Marcondes*

Departamento de Ciência da Informação, Universidade Federal Fluminense, Niterói, Brasil  
marcon@vm.uff.br



*Marília  
Alvarenga Rocha  
Mendonça*

Departamento de Ciência da Informação, Universidade Federal Fluminense, Niterói, Brasil  
marilaalvarenga@terra.com.br

*Luciana Reis Malheiros*

Departamento de Fisiologia e Farmacologia, Universidade Federal Fluminense, Niterói, Brasil  
malheiro@vm.uff.br

*Leonardo Cruz da Costa*

Departamento de Computação, Universidade Federal Fluminense, Niterói, Brasil  
leo@dcc.ic.uff.br

*Tatiana Cristina Paredes  
Santos*

Instituto Biomédico, Universidade Federal Fluminense, Niterói, Brasil  
tatianacps@biof.ufrj.br

## Resumo

Artigos científicos publicados em formato digital se constituem em bases de conhecimento científico, em especial na Medicina. Um obstáculo para o processamento semântico desse conhecimento por computadores é o fato de que, apesar do formato digital, estas bases de conhecimento são voltadas para leitura e processamento do seu conteúdo por pessoas. Propõe-se um modelo de publicação e registro para representar o conhecimento contido em artigos científicos em Medicina em formato “inteligível” por programas. Segundo o modelo, artigos científicos seriam publicados não só em formato textual, legível por pessoas, mas também como ontologias, representando o conhecimento específico contido em cada artigo. O modelo inicial foi obtido a partir de aportes teóricas de Metodologia Científica e Filosofia da Ciência e da análise de 75 artigos científicos em Medicina. O conteúdo de conhecimento científico de um artigo é associado à proposições que caracterizam fenômenos ou que estabelecem relações entre fenômenos. O modelo permite que programas “agentes de software” processem o conteúdo de cada artigo, viabilizando a recuperação semântica de informações, a avaliação da coerência, a identificação de lacunas no conhecimento científico e de novas descobertas.

## Palavras-chave

conhecimento médico; representação do conhecimento; ontologias; publicações eletrônicas; comunicação científica

Desde que foi ultrapassado o período de cultura oral com a invenção da escrita, por muitos séculos o conhecimento humano, em especial, o conhecimento científico, esteve registrado em documentos.

Estamos no limiar de uma transformação profunda nos meios de que dispõe a humanidade para registro, guarda e disseminação do conhecimento. Hoje não só podemos registrar este conhecimento em meio digital, armazenado nos dispositivos de memórias de massa, como também podemos disseminá-lo em larga escala através de redes de computadores. Um dispositivo como um *pen drive* consegue armazenar mais de 4 GB de dados, permitindo ter uma biblioteca gigantesca carregada no bolso. Mais significativo no entanto do que estas questões, é que esse conhecimento não é mais somente codificado em formato textual, legível por pessoas, não somente em formato legível por programas, que por sua vez o tornam legível por pessoas, como no caso de documentos textuais em formato WORD ou PDF, mas, o que se constitui numa real novidade, também em formatos “inteligíveis” por programas, permitindo a estes graus crescente de capacidade de realizar “inferência”, “decisões” e “raciocínios” sobre o conteúdo desses documentos. Esta é a proposta do projeto Web Semântica (Berners-Lee 2001).

Uma das bases do projeto Web Semântica são as ontologias. Uma ontologia é um modelo informacional descrevendo e representando um domínio de conhecimento específico, através dos conceitos correspondendo aos objetos relevantes nesse domínio, de sua estrutura e seus inter-relacionamentos; esse modelo deve ser de entendimento compartilhado por uma comunidade de usuários. Os conceitos são organizados em hierarquias de classes em níveis crescentes de generalidade e possuem atributos e relações entre si. Uma ontologia é representada em linguagem “inteligível” por programas “agentes de software”, e usada por estes para fazer inferências sobre os conceitos desse domínio. Quando às classes de um domínio de conhecimento específico que constituem uma ontologia são agregados representações de objetos individuais desse domínio, têm-se uma base de conhecimento. Existem muitas comunidades de usuários desenvolvendo ou utilizando ontologias, em especial na área Biomédica.

A passagem do registro de conhecimentos do formato textual para formatos suscetíveis de serem

“inteligíveis” por programas pressupõe novas formas de encará-lo. Questões óbvias e imediatas que se colocam são: o quê é o conhecimento? de quê é constituído? como pode ser sistematizado? como pode ser registrado num formato alternativa ao textual? Embora a maioria destas questões tenham sido colocadas há séculos pela Filosofia, a última questão é um problema totalmente novo e não era praticamente colocada enquanto a humanidade só conhecia a forma textual, mesmo que digital.

A pesquisa científica, em especial na área biomédica, usa de forma crescente o computador como ferramenta. Quantidades crescentes de dados sobre sequenciamento genético, proteômica etc. são mantidos em bancos de dados computacionais (Stein 2008).

O conhecimento codificado formato em “inteligível” por programas permite agenciá-los em tarefas nas quais computadores/programas são claramente mais eficientes que nós. Para o uso em larga escala da tecnologia dos “agentes de software”, deve-se extrair o conhecimento científico hoje já registrado em meio digital mas ainda em formato textual e representa-lo também num formato “inteligível” por programas.

A forma tradicional e institucionalizada através da qual a sociedade contemporânea registra e dissemina o conhecimento científico é através da publicação de artigos em periódicos. Artigos científicos se constituem em bases de conhecimento, mas para leitura e processamento por cientistas, dado ao seu formato textual. O processamento desse conhecimento, através da leitura desses artigos, sua crítica, citação, reprodução dos experimentos aí relatados, inclusão do seu conteúdo em aulas, textos didáticos, manuais e tratados, se constitui num lento processo social. Trabalhamos há anos no projeto de registrar o conteúdo de artigos científicos publicados eletronicamente em formato “inteligível” por programas. Propusemos um ambiente Web de “software” que permita a publicação eletrônica dos artigos simultaneamente de forma convencional, como texto, e em formato de ontologia, conforme ilustrado na Figura 1 a seguir. Este ambiente Web de “software” vai interagir com o autor através de um diálogo estruturado e da análise do texto do artigo, extraíndo e representando o conhecimento aí contido no formato de uma ontologia.

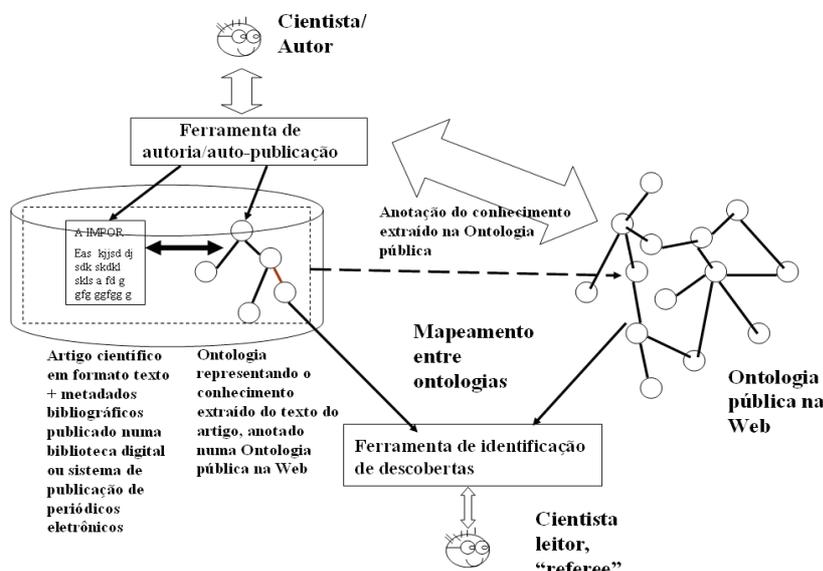


Figura 1 – Ambiente Web de autoria/autopublicação.

Criamos um modelo para os elementos semânticos que constituem o conteúdo de conhecimento de um artigo baseado nos elementos do Método Científico, na forma como eles aparecem em artigos científicos. O objetivo desse modelo é servir de base para a uma nova forma de publicar artigos em formato de ontologias, de modo a tornar seu conteúdo “inteligível” por programas, permitindo assim que esses conteúdos sejam processados de uma forma muito mais sofisticada e útil que em programas convencionais de recuperação de informações, de banco de dados, de *data mining* ou de estatística, auxiliando pesquisadores na recuperação semântica de informações, na avaliação da coerência, na identificação de lacunas no conhecimento científico e de novas descobertas.

A proposta inicial do modelo se baseava na concepção de que o conhecimento científico consiste em propor e provar a existência de *relações entre fenômenos*, até então desconhecidas. É assim que Miller (1947) define o conhecimento científico: *The above remarks imply that science is a search after internal relations between phenomena*. Um fenômeno pode ser definido como ... *an event or a process such as it appears to some human subject: it is a perceptible fact, a sensible occurrence* (Bunge 2004, p. 173).

A construção desse modelo se deu a partir de um forte referencial teórico de Filosofia da Ciência e Metodologia Científica. Esse foi complementado e validado pela análise de um conjunto de 75 artigos científicos em Ciências da Saúde, material riquíssimo que permitiu aperfeiçoar o modelo e verificar na prática as nuances de como o conhecimento científico nesta área é registrado e comunicado em artigos científicos. Este modelo foi formalizado numa Ontologia do Conteúdo de Conhecimento em artigos<sup>1</sup>. Futuros desenvolvimentos da pesquisa vão consistir em usar o modelo como base para a construção de um ambiente Web de autopublicação (Costa 2006) de artigos científicos e para o desenvolvimento de programas que comparem o conteúdo de artigos científicos registrados segundo o modelo com o conhecimento registrado em ontologias médicas, como por exemplo a UMLS – *Unified Medical Language System*<sup>2</sup> ou a *Gene Ontology*<sup>3</sup>, com vistas a identificar novas descobertas científicas (Malheiros 2005). A versão mais acabada desta proposta inicial esta descrita em Marcondes (2007).

O modelo inicial no entanto teve que ser aperfeiçoado a luz das conclusões da análise do último conjunto de artigos analisado. Trata-se de um conjunto de 15 artigos que, ao contrário dos conjuntos anteriores, guarda toda uma coerência interna e visíveis inter-relações entre si. São as chamadas *key publications*, do grupo de três pesquisadores, Elizabeth Blackburn, Carol Greider e Jack Szostak, agraciados com o Prêmio Albert Lasker de Medicina de 2006 (<http://www.laskerfoundation.org/index.html>). Os artigos cobrem o período de 1978 a 1999 e, em seu conjunto, relatam a seqüência de marcos de conhecimento da descoberta da enzima telomerase (Blackburn 2006), de seu papel fundamental na reprodução celular, sua influência em processos como

de envelhecimento celular e surgimento de câncer. Os artigos guardam forte relação uns com os outros, mostrando claramente o processo de *identificação inicial de um novo fenômeno, a agregação de novos conhecimentos de modo a caracterizá-lo cientificamente, até sua completa identificação e a investigação de suas possíveis relações com outros fenômenos*. Os resultados da análise desses artigos mostraram fatos com os quais não tínhamos nos deparado até então com a análise do material anterior e que o nosso modelo não dava conta. A análise desse conjunto de artigos impôs portanto a revisão e aperfeiçoamento do modelo inicial, de modo a torná-lo mais completo, abrangente e robusto. Este é o objetivo do presente artigo.

Ele esta estruturado assim: após esta introdução a seção 2 discute o quadro conceitual que embasa o modelo proposto; a seguir, a seção 3 discute os aspectos teóricos que embasam o modelo; a seguir, a seção 4 discute as novas características do conjunto de artigos referentes ao Prêmio Lasker 2006 que fizeram rever o modelo; a seção 5 a seguir apresenta o modelo revisto; finalmente a seção 6 apresenta conclusões e futuras direções de pesquisa.

## Material e métodos

Para a proposição do modelo foram buscados aportes teóricos de disciplinas como Ciência da Informação, em especial de Comunicação Científica, Metodologia da Ciência, Filosofia da Ciência e Ciência da Computação. Foram analisados 75 artigos em Medicina, subdivididos nos seguintes grupos: 20 artigos do periódico Memórias do Instituto Oswaldo Cruz, 20 artigos do periódico Brazilian Journal of Medical em Biological Research, ambos disponibilizados através do portal SciELO e escolhidos a partir da lista dos artigos mais consultados de ambas as publicações, disponível no site de cada uma das publicações; foram analisados ainda 20 artigos sobre células-tronco, escolhidos a partir de três importantes artigos de revisão sobre o tema.

Por fim, foram analisados 15 artigos entre as chamadas *key publications* do grupo de pesquisadores agraciados com o Prêmio Albert Lasker de Medicina do ano de 2006. O texto de Charlton (2006) indica que premiações científicas podem ajudar a identificar a elite da “ciência revolucionária”, usando as palavras de Kuhn, o que era de especial interesse para essa pesquisa.

Para este último grupo, um procedimento metodológico foi ordenar os artigos cronologicamente (ver Anexo 1) e utilizar o artigo de 2006 em que os três agraciados com o Prêmio Lasker (Blackburn 2006) comentam a trajetória das pesquisas que culminaram com a descoberta da enzima telomerase, como um guia para orientação da leitura dos 15 artigos que constituem as *key publications*.

A análise procurava identificar fenômenos descritos no artigo ou o estabelecimento de relações entre fenômenos. Após a identificação de fenômenos e relações no texto de cada artigo, verificava-se se os conceitos correspondentes aos fenômenos e relações existiam na UMLS. Os resultados da análise eram registrados em formulário específico.

A área de Medicina foi escolhida devido ao fato de que artigos científicos da área seguem um rígido padrão formal em seus textos, com seções definidas segundo o chamado padrão IMRAD – Introduction, Method, Results and Discussion -, recomendados pelo *International Committee of Medical Journals Editors*<sup>4</sup> para artigos científicos em periódicos biomédicos, facilitando assim a análise.

## Quadro conceitual

O conhecimento científico conforme veiculado através de artigos de periódicos, consistem em formular “afirmações científicas”, sobre fenômenos (Bunge 2004, p. 173) ou sobre relações entre fenômenos (Miller 1947). As afirmações científicas, conforme será visto nos exemplos da seção 4, refletem um grau crescente de certeza, apropriação e enquadramento, do fenômeno científico e de suas inter-relações, no quadro conceitual ou sistema de conceitos que compõe o conhecimento num determinado domínio científico. Este grau crescente de certeza avança na direção do que Bunge (2004, p. 3) considera ser o objetivo da Ciência, ou seja, responder as *why questions*, buscar explicações para os fenômenos.

As visões clássicas de como se faz ciência apresentadas pelos manuais de metodologia científica e pela Filosofia da Ciência em Popper (2001) e Hempel (1995), separam os procedimentos e o raciocínio empregados nas descobertas científicas da explicação metodológica dos fatos científicos. Existe toda uma ênfase nos aspectos lingüísticos, lógicos e formais da Ciência, que tem origem no positivismo lógico ainda no século XIX (Marcondes 2004).

As visões situadas dentro da “lógica da justificativa”, sejam elas dos positivistas lógicos ou de seus críticos como Popper, são por demais formais e os exemplos analisados mostram que, numa área de pesquisa de ponta atual como a biologia celular, fornecem poucos subsídios para se compreender as práticas de pesquisa e descoberta que levaram à construções científicas que resultaram na descoberta da telomerase. Assim, o material por nós analisado, conforme será mostrado na seção 4, apresenta bastante semelhança com as colocações de Aliseda (2004), Klahr e Simon (1999) e Thagard (1993).

A distinção entre “lógica da descoberta”, em oposição à “lógica da justificativa” é enfatizado em Atocha (2004), numa crítica às visões positivistas que afirmam que: “The context of discovery is taken to be purely psychological” (Aliseda 2004, p. 340).

Ou então:

*Reinchenbach's distinction between the contexts of justification and of discovery has left out of its analysis – especially from a formal point of view – a very important part of scientific practice, that which includes issues related to the generation of new theories and scientific explanations, concept formation as well as aspects of progress and discovery in science* (Aliseda 2004, p. 341).

Thagard, ao criticar tanto o método indutivo, privilegiado pelos positivistas lógicos, quanto o método

hipotético-dedutivo de Popper, chamando a ambos de mitos (Thagard 1993, p.176), afirma ao contrário que:

*In well-trod areas of investigation, it may be possible to form a Sharp hypothesis and then test it. But when novel topics are being pursued, researchers in psychology and other fields cannot always start with hypotheses sharp enough to be tested. Often some vague ideas will lead to the collection of some data, which then suggested a refinement of an existing hypothesis. Or results are very different from what was expected may spur abductive formation of a new hypothesis that can then be subject to further test* (Thagard 1993, p. 177).

Klahr e Simon (1999, p.8) também criticam Popper e dão especial ênfase ao contexto da descoberta:

In science there is an important, and extremely common, form of experiment, at times referred to somewhat dismissively as “exploratory,” that is guided by no specific hypothesis to be tested, and no clear control condition, but only a vague and general direction of inquiry. The goal of exploratory experiments is to permit phenomena to appear that will invite exploration or suggest whole new forms of representation or generate new hypotheses.

*The contemporary literature on research methodology is dominated by the notion, promulgated by Popper (1959) among others, that the purpose of observation in general, and experiment in particular, is to test hypotheses in order either to falsify or validate them. In contrast to this position, we have argued that much of the important empirical work in science is undertaken – to use Reichenbach's phrase – in the context of discovery rather than the context of verification (see Simon 1973). That is, a major goal of empirical work in science is to discover new phenomena and generate hypotheses for describing and explaining them, and not simply to test hypotheses that have already been generated. Indeed, theories cannot be tested until they have been created, and creation takes place in the context of discovery, not verification.*

Estamos interessados em buscar um modelo para representar o conhecimento contido no texto de artigos científicos em formato “inteligível” por programas, permitindo a estes programas graus crescente de “inferência”, “decisões” e “raciocínios” sobre o conteúdo desses artigos.

Historicamente a intervenção da humanidade no seu ambiente é cada vez mais indireta, através de diferentes ferramentas. Em especial nas ciências as ferramentas usadas são ferramentas cognitivas. Representar formalmente o conhecimento científico em geral e nas Ciências da Saúde em particular em formato “inteligível” por programas significa em termos práticos desenvolver ontologias. Essas ontologias serão tão mais úteis quanto mais acuradamente refletirem as descobertas científica nesta área, ou seja, quanto mais corresponderem ou forem análogas à realidade desta área, poderem ser usadas como *modelos* científicos desta realidade, modelos esses passíveis de serem processáveis por computadores para testarem hipóteses, fazerem comparações, diagnósticos, identificarem inconsistências etc., se constituindo em ferramentas cognitivas e instrumentos para o avanço da pesquisa e do conhecimento científico.

O termo inferência em Lógica significa o processo (também chamado de “raciocínio”) de derivar conse-

quências verdadeiras de premissas verdadeiras ou tidas como verdadeiras. Simular processos de inferência num ambiente computacional seria o processo em que, a partir de informações fornecidas a um sistema, este retornaria outras informações que estejam de alguma maneira relacionadas às informações fornecidas.

O conhecimento científico conforme veiculado através de artigos de periódicos, consiste em formular, através da linguagem, proposições contendo *afirmações científicas* sobre fenômenos (Bunge 2004, p. 173), ou relacionando fenômenos entre si (Miller 1947) ou relacionando um fenômeno a suas características.

Buscamos conceituações de “fenômeno” que possam servir para representar formalmente o conhecimento científico em Medicina, objeto dessa pesquisa. Uma definição de fenômeno usada em textos de Filosofia e Metodologia da Ciência seria: ... *an event or a process such as it appears to some human subject: it is a perceptible fact, a sensible occurrence* (Bunge 2004, p. 173).

Afirmarões científicas refletem um grau crescente de certeza, apropriação do fenômeno científico e de suas inter-relações, no quadro conceitual ou sistema de conceitos que compõe o conhecimento num determinado domínio científico.

Foram identificadas na análise da literatura que constitui o material empírico duas formas de conhecimento científico enquanto relações:

A primeira forma seria a apropriação de um fenômeno através da progressiva caracterização do mesmo através coleta sistemática de “afirmações científicas” (Bunge 2004, p. 173) sob a forma de proposições *relacionando o fenômeno à suas características* (Dalhberg 1977, p. 16). Mais especificamente, Dahlberg chama de “características essenciais” aquelas que caracterizam ou dão identidade a um determinado fenômeno e que sem as quais esse fenômeno perderia sua identidade (Guarino 1997). São características que constituem o que Aristóteles chama de Essência ou atributos essenciais da substância (Chauí 2005). Através da coleta sistemática de suas características, testadas cientificamente, um fenômeno é progressivamente identificado e integrado ao sistema de conceitos de um domínio científico.

A segunda forma de conhecimento científico seria a *identificação e o estabelecimento de relações entre fenômenos distintos*, até então desconhecidas.

Relações, na forma proposta acima, também refletiriam graus crescentes de certeza das proposições científicas, desde uma Questão ou Problema: qual o mecanismo que determina a síntese das extremidade dos telômeros? Onde um dos *relata* é desconhecido, passando por uma Hipótese, onde a relação entre os *relata* é hipotética, uma atividade enzimática determina a síntese das extremidade dos telômeros? Até uma Conclusão: a enzima telomerase determina a síntese das extremidade dos telômeros, onde a relação entre os *relata* é provada por um experimento.

No entanto a palavra fenômeno está eivada de conotações subjetivas, necessariamente incompatíveis com o conhecimento científico, por sua vinculação à

Fenomenologia (Chauí 2005). Esta disciplina da filosofia estuda os fenômenos como percebidos por um observador individual, em oposição a real natureza das coisas, o ser real, ou seja, as aparências em oposição à realidade.

Se a observação de um fenômeno pode ser distorcida pelo observador, como já têm sido largamente discutido, se este esta condicionado socialmente e historicamente e, em termos científicos, como já mostrou Kuhn, paradigmaticamente, se o conhecimento é construído pelo indivíduo progressivamente, como diz Piaget (1978), o conhecimento científico é uma construção eminentemente social (Ziman 1979). A Ciência enquanto instituição tem mecanismos que asseguram um alto grau de consenso para um determinado estágio de conhecimento num determinado momento histórico. É claro que este estágio de conhecimento provisório, superável, limitado foi (socialmente) construído. No entanto, corresponde ao que é consensualmente identificado como *correspondendo* à realidade, ao estágio de conhecimento sobre a realidade. Na medida em que este conhecimento é partilhado e consensado socialmente, na medida que, com base nele, podemos intervir na realidade, usá-lo como ferramenta cognitiva fazendo previsões sobre ela, este conhecimento – uma representação mental ou um registro, inscrição ou documento capaz de ser apropriada intersubjectivamente – *corresponde à realidade*. Assim, a pesquisa científica, ao observar e estudar os fenômenos, fornece os elementos - o conhecimento científico – para a construção de uma sempre provisória, sempre em construção, inacabada, Ontologia. Esta deve corresponder, o máximo possível num determinado estágio de conhecimento de uma dada ciência, à realidade mesmo.

Barry Smith (2002, p. 2), discutindo a relação entre Ciência e Ontologia (enquanto domínio de conhecimento preocupado com a natureza dos seres), afirma que Ontologia não pretende “explicar” a natureza como a Ciência, seu papel seria vir a seguir das explicações para descrever, organizar e sistematizar o conhecimento obtido pelas descobertas científicas. Este parece ser um lugar a ser ocupado também pela Ciência da Informação.

## Resultados

O processo de crescente caracterização e apropriação científica de um fenômeno e a posterior identificação de relações entre esse fenômeno e outros é ilustrado nos seguintes exemplos e pode ser acompanhado pelos títulos dos artigos pertencentes ao grupo da premiação Lasker 2006, ordenados cronologicamente no Anexo 1:

- No artigo mais antigo analisado do grupo referente ao Prêmio Lasker 2006 (Blackburn & Gall 1978), vê-se esse aspecto da gradual caracterização de um novo fenômeno científico, não só no título, mas também nos objetivos do artigo, colocados no seu abstract: *The extrachromosomal genes coding for the ribosomal RNA in the ciliated protozoan Tetrahymena thermophila we studied with respect to sequences occurring at their termini* (Blackburn 1998, p. 33).

- Num artigo de revisão que mostra um quadro atual da pesquisa sobre telomerase, Cech (2004) afirma que

o objetivo da pesquisa que conduziu a descoberta da telomerase seria identificar a *entidade* (termo por sinal, bastante usado na modelagem de ontologias) responsável pela replicação das extremidades dos cromossomos, até então desconhecida, não caracterizada, não integrada ao quadro do conhecimento existente até então: *Carol Greider, a graduate student in Liz Blackburn's group at the University of California, Berkeley, had chosen an ambitious PhD thesis project: identify the molecular entity responsible for replicating chromosome ends* (Cech 2004, p. 273).

- No mesmo artigo Cech (2004) esclarece que o objetivo das pesquisas de Greider e Blackburn no artigo que marca a descoberta da telomerase seria: *The identification and characterization of this new enzymatic activity was the subject of Greider and Blackburn* (1985).

- Outro exemplo retirado de um artigo do mesmo grupo, mostra que a telomerase ainda não estava identificada como uma enzima (*a terminal transferase-like activity*) nem estava clara a sua relação com o processo de complementação dos telômeros (*which adds the host cell telomeric sequence repeats onto recognizable telomeric ends*): *Based on all these considerations, the proposal was made that telomere replication involves a terminal transferase-like activity which adds the host cell telomeric sequence repeats onto recognizable telomeric ends.* (Shampay et al. 1984), citado em Greider (1985, p. 405).

- Num artigo subsequente do mesmo conjunto, verifica-se uma crescente apropriação científica do mesmo fenômeno. É identificada uma atividade enzimática de transferase, explicitada no título do artigo *Identification of a specific telomere terminal transferase activity in Tetrahymena extracts* (Greider CW, Blackburn EH. Cell. 1985; 43: 405-413).

- No mesmo artigo, as autoras propõem, ainda com pouco grau de certeza, o enquadramento da atividade da telomerase dentro do quadro conceitual já conhecido: *The authors made the reasonable proposal that the activity might be related to known terminal transferases, such as the enzyme that adds CCA to the 3' ends of transfer RNAs* (Cech 2004, p. 273).

- Posteriormente esta atividade enzimática é identificada, ou seja, enquadrada no sistema de conceitos desse domínio científico específico, ou seja, numa classificação de substâncias, e essa enzima é finalmente batizada de telomerase com este nome:

Greider CW, Blackburn EH. *The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity.* Cell. 1987; 51: 887-898.

- Posteriormente a atividade, função ou papel da telomerase como catalisador e molde no processo de sínteses e complementação das extremidades dos telômeros é identificada:

Greider CW, Blackburn EH *A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis.* Nature. 1989; 337: 331-337.

- Ou então: *Our results indicate the involvement of such sequence-specific telomeric DNA-protein interaction in cell or nuclear division* (Yu 1990, p. 131).

Uma vez caracterizado o fenômeno da complementação das extremidades dos telômeros pela ação da enzima telomerase, começam a surgir propostas estabelecendo relações desse fenômeno com dois outros: a senescência celular, ou seja, o número finito de vezes que uma célula é capaz de se reproduzir, consequência do encurtamento progressivo dos telômeros a cada duplicação e da incapacidade da célula em complementá-los através da ação da telomerase, que conduz a morte celular; e a relação da telomerase com o câncer, identificado como um processo descontrolado de duplicação celular. Esses casos são descritos a seguir:

- *These mutations also lead to nuclear and cell division defects, and senescence, establishing an essential role for telomerase in vivo* (Yu 1990, p. 126).

Ou então, no artigo

- Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, Futcher CW, Greider CW, Harley CB. *Telomere length predicts the replicative capacity of human fibroblasts.* Proc. Natl. Acad. Sci. USA. 1992; 89: 10114-10118.

Ou nos trechos, e no artigo seguinte: *This shortening has been proposed to play a role in signaling the cell cycle exit characteristics of senescent cells (14, 15), although a causal role has not been demonstrated* (Prowse 1993, p. 1493). Nesse artigo os autores propõem a existência do que é chamado aqui de uma relação “fraca” ao afirmar que uma relação causal ainda não foi demonstrada. No outro trecho e no artigo é proposta uma relação entre a atividade da telomerase e o câncer: *It has been proposed that the finite cell division capacity of human somatic cells is limited by telomere length (10). This is consistent with reports that telomerase activity is often high in cancer and immortalized tissue culture cells* (Mceachern 1995, p. 403).

E ainda no artigo:

- Rudolph KL, Chang S, Lee HW, Blasco M, Gottlieb G, Greider CW, DePinho RA. *Longevity, stress response, and cancer in aging telomerase deficient mice.* Cell. 1999; 96: 701-716.

No artigo mais antigo analisado do grupo referente ao Prêmio Lasker 2006 (Blackburn & Gall 1978), vemos esse aspecto da gradual caracterização de um novo fenômeno científico, não só no título, mas também nos objetivos do artigo, colocados no seu abstract: *The extrachromosomal genes coding for the ribosomal RNA in the ciliated protozoan Tetrahymena thermophila we studied with respect to sequences occurring at their termini* (Blackburn 1998, p. 33).

## Um modelo para conhecimento médico em artigos científicos

A seguir um modelo para representar, em formato “inteligível” por programas o conhecimento contido em artigos científicos em Medicina é proposto. São mostradas as bases do modelo, a partir do conteúdo dos artigos que constituem o campo empírico e de sua análise dentro do quadro conceitual descrito anteriormente. Como já foi mencionado, a versão anterior do modelo (Marcondes

2007) enfatizava o papel das hipóteses enquanto relações entre fenômenos, segundo uma visão mais convencional da Ciência, baseada no Método Hipotético-dedutivo; a versão atual do modelo, apresentada aqui, incorpora elementos que representam o conhecimento científico sendo progressivamente construído, através da caracterização e incorporação de um novo fenômeno.

Artigos científicos distinguem-se pelo tipo de raciocínio que empregam ao conduzirem a argumentação acerca dos fenômenos discutidos. Existem artigos teóricos e artigos experimentais. Esta classificação é baseado em Hutchins (1997) e Gross (1990) e em textos a partir da visão de abdução em Pierce (1977), como processo de descoberta de novos *insights* em Ciência (Hoffman 1997, Magnani 2001, Paavola 2004, Aliseda 2004).

**Artigos teórico-abdutivos** se caracterizam por discutirem questões de maior abrangência. Analisam criticamente diversas hipóteses anteriores, mostrando suas fragilidades. Estes artigos são os que têm mais potencial de apresentarem contribuições para a Ciência, já que discutem ou questionam o paradigma vigente (Kuhn 2003). Sua contribuição é uma nova hipótese, indicando um novo caminho de pesquisa. O tipo de raciocínio empregado é o abdução ou seja, o *insight* sobre a solução de questões não explicadas na Ciência e a formulação de novas hipóteses de solução-las.

**Artigos experimentais** constam necessariamente de um experimento empírico; dividem-se em exploratórios, dedutivos e indutivos. Caracterizam-se por discutirem questões num escopo de abrangência limitado. Não discutem os rumos de uma teoria científica, mas se limitam a confirmá-la ou aperfeiçoá-la. Sempre trazem resultados experimentais.

**Artigos experimentais-exploratórios** têm um caráter exploratório ao desvendar e buscar caracterizar um fenômeno, trabalhando na direção proposta por Dahlberg de formular e provar proposições que caracterizam um fenômeno,

**Artigos experimentais-dedutivos** trabalham a partir de relações entre fenômenos já formuladas anteriormente, cujas referências vêm citadas, aplicando-as a testando-as e validando-as um contexto específico. Os **artigos experimentais-indutivos** se caracterizam por proporem e testarem novas relações entre fenômenos.

A estrutura textual dos artigos em Ciências da Saúde segue o padrão IMRAD, como já foi mencionado. Essa estrutura corresponde à “surface structure” de Chomsky (1981) e a microestrutura de Kintsh e Van Dijk (1972). Já os componentes semânticos num artigo, que compõe o modelo proposto, correspondem à *deep structure* de Chomsky e à macroestrutura de Kintsh e Van Dijk; são descritos a seguir, identificados em maiúsculas.

Um PROBLEMA expressa uma carência, insatisfação ou deficiência conceitual com o atual estado de conhecimento num domínio. Um PROBLEMA pode se desdobrar em OBJETIVOS de pesquisa e, eventualmente, na formulação mais precisa de uma QUESTÃO que endereça a deficiência conceitual; esta QUESTÃO pode ser referir a um FENÔMENO (nos artigos EX-

PLORATÓRIOS), ou a dois ou mais FENÔMENOS envolvidos numa RELAÇÃO\_ENTRE\_FENÔMENOS ou HIPÓTESE. Uma HIPÓTESE relaciona dois ou mais FENÔMENOS através de um TIPO-DE-RELAÇÃO.

Um autor num artigo pode formular uma hipótese original – HIPÓTESE(o) ou tomar a hipótese prévia – HIPÓTESE(p) - de outros autores; neste caso uma ou mais citações referentes à HIPÓTESE(p) – CITAÇÕES(h) - são feitas. Um autor também pode analisar várias HIPÓTESES(p) para mostrar que elas são insatisfatórias como soluções para o PROBLEMA e formular sua HIPÓTESE(o). Um artigo teórico se justifica simplesmente por propor uma nova HIPÓTESE(o).

Da hipótese, num artigo experimental, deve ser derivado um EXPERIMENTO capaz de ser observável empiricamente. Em um artigo científico EXPERIMENTAL, significa ter RESULTADOS observados segundo determinada MEDIDA, em determinado CONTEXTO segundo determinada METODOLOGIA. Este CONTEXTO onde os FENÔMENO(s) relacionados na HIPÓTESE são observados pode ser desdobrado em AMBIENTE – comunidade ou instituição onde o fenômeno ocorre -, ESPAÇO - o lugar onde o fenômeno ocorre -, TEMPO ou época em que o fenômeno ocorre e GRUPO de indivíduos onde o fenômeno ocorre. Todo artigo também traz uma CONCLUSÃO, na forma de uma proposição sobre um fenômeno ou sobre RELAÇÕES\_ENTRE\_FENÔMENOS.

O desenvolvimento do raciocínio num **artigo teórico abdução** segue o seguinte padrão:

- dado um PROBLEMA, com os seguintes aspectos e dados

- os seguintes Autores/HIPÓTESES anteriores para sua solução não são satisfatórias,

- *diante disso, propomos a seguinte HIPÓTESE original.*

O desenvolvimento do raciocínio num **artigo experimental dedutivo** segue o seguinte padrão:

- dado um PROBLEMA, com os seguintes aspectos e dados,

- os seguinte Autores formularam HIPÓTESES anteriores para sua solução,

- *diante disso, escolhemos a seguinte (uma das HIPÓTESE anteriores).*

Ampliamos e re-contextualizamos esta HIPÓTESE anterior; desenvolvemos o seguinte EXPERIMENTO para testar esta HIPÓTESE anterior;

- o EXPERIMENTO apresentou os seguintes RESULTADOS.

O desenvolvimento do raciocínio num **artigo experimental indutivo** segue o seguinte padrão:

- dado um PROBLEMA, com os seguintes aspectos e dados,

- uma solução para este PROBLEMA pode se basear na seguinte HIPÓTESE,

- desenvolvemos o seguinte EXPERIMENTO para estar esta HIPÓTESE,

- estes testes apresentaram os seguintes RESULTADOS.

O desenvolvimento do raciocínio num artigo **experimental exploratório** segue o seguinte padrão:

- dado um PROBLEMA ou FENÔMENO ainda não bem caracterizado,

- desenvolvemos o seguinte EXPERIMENTO que permite identificar a(s) seguinte(s) CARACTERÍSTICA(S) desse FENÔMENO.

Estes esquemas resultaram no atual modelo ou **Ontologia do Conteúdo de Conhecimento em Artigos Científicos – OCCAC**, também ilustrado na Figura 2:

Classes: artigos TEÓRICOS

têm componentes

PROBLEMA  
HIPÓTESE(a)  
HIPÓTESE(o)  
CONCLUSÃO(ões) e

artigos EXPERIMENTAIS

Subclasses: artigos EXPLORATÓRIOS

têm componentes

PROBLEMA  
FENÔMENO  
EXPERIMENTO  
CONCLUSÃO(ões)

artigos INDUTIVOS

têm componente

PROBLEMA  
HIPÓTESE(o)

EXPERIMENTO  
CONCLUSÃO(ões) e

artigos DEDUTIVOS

têm componente

PROBLEMA  
HIPÓTESE(a)  
EXPERIMENTO  
CONCLUSÃO(ões)

COMPONENTES semânticos de artigos

PROBLEMA  
Subcomponentes: OBJETIVOS  
QUESTÃO de pesquisa

HIPÓTESE (prévia ou nova)  
Subcomponentes: FENÔMENO(S)  
TIPO-DE-RELAÇÃO

REFERENCIAS (somente nas HIPÓTESES prévias)

FENÔMENO (um, nos artigos EXPLORATÓRIO(s))  
Subcomponentes: CARACTERÍSTICAS(s)

EXPERIMENTO

Subcomponentes: RESULTADOS (dados quantitativos)

MEDIDA

CONTEXTO

Subcomponentes:

ESPAÇO

TEMPO

GRUPO social

CONCLUSÃO(ões)

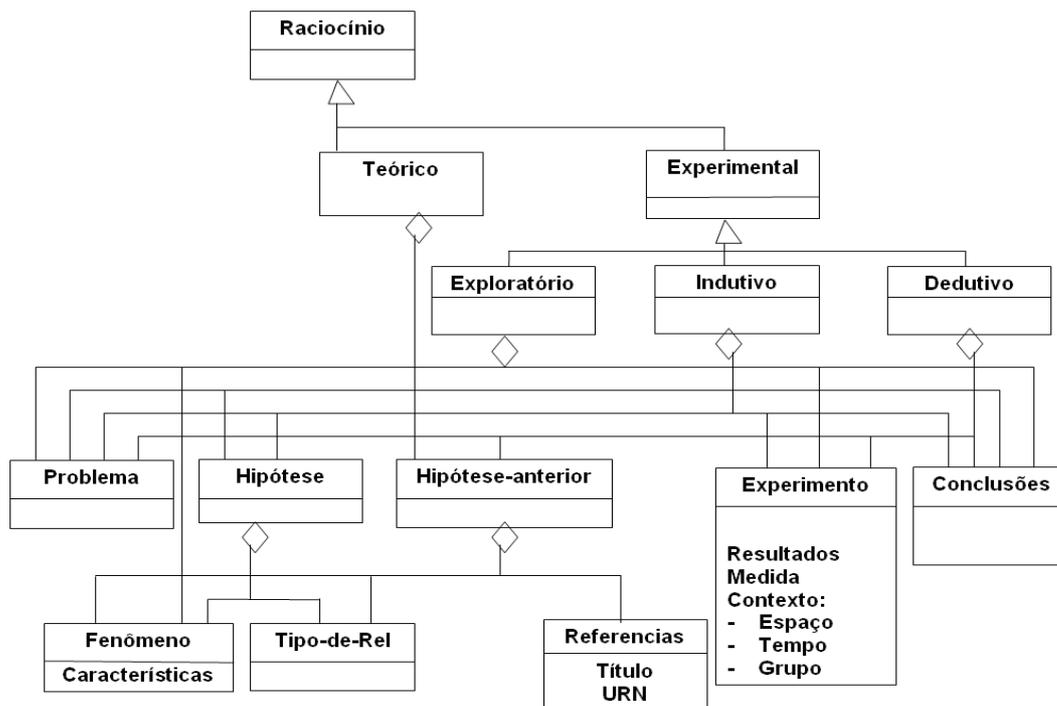


Figura 2 - Ontologia do Conteúdo de Conhecimento em Artigos Científicos Digitais.

A análise do seguinte artigo mostra como componentes semânticos de um artigo (neste caso a hipótese) seriam identificados e registrados segundo o modelo.

- Camara GNI, Cerqueira DM, Oliveira APG et al. *Prevalence of human papillomavirus types in women with pre-neoplastic and neoplastic cervical lesions in the Federal District of Brazil*. Mem. Inst. Oswaldo Cruz. [online]. 2003 Oct.; 98:7.

3 passos:

Passo 1 – Tipo de raciocínio identificado: experimental-dedutivo, ou seja, o artigo realiza um experimento para mostrar a prevalência de HPV, hipótese já formulada anteriormente por outro autor.

Passo 2 – Elementos semânticos de conhecimento são identificados no texto, como a hipótese formulada pelo autor:

Hipótese (anterior)

Antecedente: HPV

Tipo de Relação: causa

Conseqüente: lesões cervicais pre-neoplásticas e neoplásticas

Passo 3 – Cada um desses elementos é mapeado em termos ou relações da base de conhecimento pública, a UMLS, UMLS Semantic Network

Papillomavirus, Human

“Causes”, UMLS Semantic network relação R147

Colonic Neoplasms

## Conclusões

A Ciência da Informação vêm de uma longa tradição teórica, metodológica e prática que converge para muitas das questões atuais colocadas pela proposta da Web semântica e para a construção de ontologias. Uma área de pesquisa recorrente na Ciência da Informação é o processamento semântico de informações por computadores, com contribuições já históricas como as de Shera (1957), Luhn (1960) e mais recentes como as de Gardin (2001) e Kajikawa (2006).

A Ciência da Informação pode e deve ir além de prover técnicas e metodologias para permitir acesso ao texto completo de artigos científicos em bibliotecas digitais e repositórios científicos. As metodologias atuais de indexação consistem em assinalar a registros bibliográficos palavras-chave ou termos de um vocabulário controlado isolados, sem nenhuma relação ou papel semântico entre si. Mas as pesquisas em CI relativas à importância das relações e seu papel semântico também já são históricas, como atesta a recente revisão de Khoo (2007).

O modelo proposto, ao identificar tipos de relações (com sua semântica própria) e características dos fenômenos descritos num artigo, permite realizar inferências automáticas e, por exemplo, resolver consultas sofisticadas como:

- que (outros) artigos (também) têm hipóteses relacionando HPV como causa de lesões pré-neoplásticas e neoplásticas em mulheres?

- que artigos têm hipóteses relacionando outros fatores que não HPV como causa de lesões pré-neoplásticas e neoplásticas em mulheres?

- que artigos identificam outras características relacionadas com a estrutura das extremidades das moléculas lineares de rDNA?

- que artigos identificam características do fenômeno de replicação de telômeros que podem ser associadas à atividade enzimática?

A implementação prática do modelo de registro do conteúdo de artigos científicos conforme descrito pressupõe o desenvolvimento de todo um conjunto de ferramentas de software que processem conteúdos estruturados segundo o proposto. Trata-se na verdade de um programa de pesquisa. Acreditamos que um ponto de partida sólido para isto é estabelecer o modelo proposto aqui. Duas destas aplicações inicialmente visualizadas estão descritas em Malheiros (2005) e Costa (2006), e delineadas na Figura 1.

O modelo mostra os benefícios de um formato semanticamente rico para o registro do conteúdo de artigos científicos, baseado em relações, viabilizando que, através delas, programas agentes de “software” realizem “inferências”. Forsythe (1989), falando sobre as primeiras experiências de construção de sistemas especialistas nas décadas de 1970 e 1980, chama a aquisição de conhecimento de *o gargalo da construção de sistemas especialistas*. Realizar a *aquisição de conhecimento* diretamente a partir dos autores/pesquisadores a partir de seus artigos científicos, que já trazem um alto grau de formalização desse conhecimento, como é a presente proposta, pode se mostrar uma alternativa promissora. Com o apoio de ferramentas de software adequadas, o conteúdo de artigos científicos pode ser extraído como um subproduto do processo de autoria/autopublicação de artigos científicos pelos próprios autores, hoje já bastante comum quando estes submetem seus trabalhos a repositórios, publicações eletrônicas ou bibliotecas digitais. A visão delineada aqui (Figura 1) resultará em publicações eletrônicas com um potencial de tratamento semântico de seu conteúdo por programas agentes de “software” muito mais rico do que é possível nas publicações eletrônicas textuais atuais.

## Notas

1. Disponível em [www.professores.uff.br/marcondes/Scientific\\_reasoning.owl](http://www.professores.uff.br/marcondes/Scientific_reasoning.owl)

2. Disponível em [www.nlm.nih.gov/pubs/factsheet/umls.html](http://www.nlm.nih.gov/pubs/factsheet/umls.html)

3. Disponível em <http://www.geneontology.org/>

4. Disponível em [www.icmje.org](http://www.icmje.org)

## Referências bibliográficas

Aliseda A. Logics in scientific discovery. *Foundations of Science*. 2004; 9: 339-63.

Berners-Lee T, Hendler J, Lassila O. The semantic web. *Sci Am*. 2001 May. Disponível em <<http://www.scian.com>>

com/2001/0501issue/0501berners-lee.html>. Acesso em 24 maio 2001.

Blackburn EH, Greider CW, Szostak J. Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nature*. 2006 October; 12(10).

Bunge M. *Philosophy of Science*. New Brunswick, London: Transaction Publishers; 1998.

Cech TR. Beginning to understand the end of the chromosome. *Cell*. 2004 Jan.; 116:273-9.

Charlton BG. Editorial. Scientometric identification of the elite 'revolutionary science' research institutions by analysis of trends in Nobel prizes 1947-2006. *Medical Hypotheses*. 2007; 68:931-4.

Chauí M. *Convite à Filosofia*. São Paulo: Érica; 2005.

Chomsky N. *Regras e representações: a inteligência humana e seus produtos*. Rio de Janeiro: Ed. Zahar; 1981.

Costa LC da. Uma ferramenta para edição, extração e representação do conhecimento contido em artigos científicos publicados na Web. Projeto de Tese de Doutorado para ingresso no PPGCI UFF/IBICT. Niterói; 2006.

Dahlberg I. *Ontical structures and universal classification*. Bangalore: Sarada Ranganathan Endowment for Library Science; 1978.

Forsythe DE, Buchanan BG. Knowledge acquisition for expert systems: some pitfalls and suggestions. *IEEE Transactions on Systems, Man and Cybernetics*. 1989 May/June; 19(3):435-42. Disponível em: <<http://ieeexplore.ieee.org/iel1/21/1336/00031050.pdf?tp=&isnumber=&arnumber=31050>>. Acesso em: 21 abr. 2008.

Gardin J-C. Vers un remodelage des publications savantes: ses rapports avec sciences de l'information. In: *Colloques Isko-France. Filtrage et résumé automatique de l'information sur les réseaux*, Conference invitee, Université de Nanterre – Paris X; 2001. (Conference proceedings).

Gross AG. *The Rhetoric of Science*. Cambridge, EUA; London, Inglaterra: Harvard University Press; 1990.

Guarino N. Some organizing principles for a unified top-level ontology. New version of paper presented at AAAI Spring Symposium on Ontological Engineering, Stanford University; March 1997.

Hempel K. *Aspects of scientific explanation: and other essays in the philosophy of science*. New York: Free Press; 1965.

Hutchins J. On the structure of scientific texts. In: *UEA Papers in Linguistics*, 5 th., 1977, Norwich. **Proceedings**. Norwich, UK: University of East Anglia, 1977. p. 18-39. Disponível em: <<http://ourworld.compuserve.com/homepages/wjhutchins/UEAP/L-1977.pdf>>. Acesso em: 30 mar. 2006.

Kajikawa Y, Abe K, Noda S. Filling the gap between researchers studying different materials and different methods: a proposal for structured keywords. *J Inform Sci*. 2006; 32: 511-24.

Khoo C, Na JC. Semantic Relations in Information Science. *Ann Rev Inform Sci Technol*. 2007; 157-228.

Kintsch W, Van Dijk TA. Towards a model of text comprehension and production. *Psychol Rev*. 1972; 84(5):363-93.

Klahr D, Simon HA. *Studies of scientific Discovery: complementary approaches and convergent findings*. *Psychol Bull*. 1999; 125(5): 524-43.

Kuhn TS. *A estrutura das revoluções científicas*. São Paulo: Perspectiva; 2003. (Série Debates Ciência).

Luhn H. Keyword in Context Index for Technical Literature. *American Documentation*. 1960; 11(4): 288-95.

Malheiros L. A identificação de novas descobertas científicas através da análise do conhecimento contido em artigos científicos. Projeto de Tese de Doutorado para ingresso no PPGCI UFF/IBICT. Niterói; 2005.

Marcondes D. *Filosofia analítica*. Rio de Janeiro: Jorge Zahar; 2004. (Coleção Passo a passo).

Marcondes CH, Mendonça MAR, Malheiros CLC da, Santos TCP, Pereira LG. Representing and coding the knowledge embedded in texts of Health Science Web published articles. In: Chan L, Marten B. Ed. *ICCC EIPub - International Conference on Electronic Publishing*, Viena, Austria; 2007. Disponível em <<http://elpub.scix.net>>.

Miller DL. Explanation versus description. *Philosophy Rev*. 1947; 56(3): 306-12.

Piaget J. *Psicologia e Epistemologia: por uma teoria do conhecimento*. Rio de Janeiro: Forense; 1978.

Popper K. *A lógica da pesquisa científica*. São Paulo: Ed. Cultrix, Ed. USP; 2001.

Shera JH, Kent A, Perry JW, editors. *Information systems in documentation*. New York: Interscience Publishers; 1957. (Advances in Documentation and Library Science, v. 1).

Smith B. Beyond concepts: ontology as reality representation. In: *FOIS - International Conference on Formal Ontology and Information Systems*, Turin; nov. 2004. Disponível em: <http://ontology.buffalo.edu/bfo/BeyondConcepts.pdf>. Acesso em: 16 mar. 2007.

Smith B. *Ontology and information systems*. 2002. Disponível em: <[http://ontology.buffalo.edu/ontology\(PIC\).pdf](http://ontology.buffalo.edu/ontology(PIC).pdf)>. Acesso em: 26 maio 2008.

Stein LD. Towards a cyberinfrastructure for the biological sciences: progress, visions and challenges. *Nature Rev Gen*. 2008 Sept.; 9:678-88.

Thagard P. *Computational Philosophy of Science*. Cambridge: The MIT Press; 1993.

Ziman J. *Conhecimento público*. Belo Horizonte: Itatiaia, São Paulo: Ed. da Universidade de São Paulo; 1979.

## Anexo 1 – “Key publications” do Prêmio Lasker/2006, ordenados cronologicamente

Ano	Artigo
1978	Blackburn, E.H. and Gall, J.G. (1978) A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in <i>Tetrahymena</i> . <i>J. Mol. Biol.</i> 120: 33-53.
1987	Szostak, J.W. and Blackburn, E.H. (1982) Cloning yeast telomeres on linear plasmid vectors. <i>Cell</i> 29: 245-255.
1983	Murray, A.W. and Szostak, J.W. (1983) Construction of artificial chromosomes in yeast. <i>Nature</i> 305: 189-193.
1984 JAN	Shampay, J., Szostak, J.W., and Blackburn, E.H. (1984) DNA sequences of telomeres maintained in yeast. <i>Nature</i> 310: 154-157.
1984 MAIO	Dunn, B.L., Szauter, P., Pardue, M-L., Szostak, J.W. (1984) Transfer of telomere-adjacent sequences to linear plasmids by recombination. <i>Cell</i> 39: 191-201.
1985	Greider, C.W. and Blackburn, E.H. (1985) Identification of a specific telomere terminal transferase activity in <i>Tetrahymena</i> extracts. <i>Cell</i> 43: 405-413.
1987	Greider, C.W. and Blackburn, E.H. (1987) The telomere terminal transferase of <i>Tetrahymena</i> is a ribonucleoprotein enzyme with two kinds of primer specificity. <i>Cell</i> 51: 887-898.
1988 NOV	Greider, C.W. and Blackburn, E.H. (1987) A telomeric sequence in the RNA of <i>Tetrahymena</i> telomerase required for telomere repeat synthesis. <i>Nature</i> 337: 331-337.
1989 JAN	Lundblad V. and Szostak, J.W. (1989) A mutant with a defect in telomere maintenance leads to senescence in yeast. <i>Cell</i> 57: 633-643.
1990	Yu, G.L., Bradley, J.D., Attardi, L.D. and Blackburn, E.H. (1990) In vivo alteration of telomere sequences and senescence caused by mutated <i>Tetrahymena</i> telomerase RNAs. <i>Nature</i> 344: 126-132.
1992	Allsopp, R.C., Vaziri, H., Patterson, C., Goldstein, S., Younglai, E.V., Futcher, C.W., Greider, C.W., and Harley, C.B. (1992) Telomere length predicts the replicative capacity of human fibroblasts. <i>Proc. Natl. Acad. Sci. USA</i> 89: 10114-10118.
1993	Prowse, K.R., Avilion, A.A., and Greider, C.W. (1993) Identification of a nonprocessive telomerase activity from mouse cells. <i>Proc. Natl. Acad. Sci. USA</i> 90: 1493-1497.
1995	McEachern, M.J. and Blackburn, E.H. (1995) Runaway telomere elongation cause by telomerase RNA mutations. <i>Nature</i> 376: 403-409.
1999	Rudolph, K.L., Chang, S, Lee, H.W., Blasco, M., Gottlieb, G., Greider, C.W., and DePinho, R.A. (1999) Longevity, stress response, and cancer in aging telomerase deficient mice. <i>Cell</i> 96: 701-716
2001	Kim, M.M., Rivera, M.A., Botchkina, I.L, Shalaby, R., Thor, A.D., and Blackburn, E.H. (2001) A low threshold level of expression of mutant-template telomerase RNA is sufficient to inhibit tumor cell growth. <i>Proc. Natl. Acad. Sci. USA</i> 98: 7982-7987

## Sobre os autores

### *Carlos Henrique Marcondes*

<http://www.professores.uff.br/marcondes>.

Professor do Departamento de Ciência da Informação da UFF, onde atua nos cursos de graduação em Arquivologia e Biblioteconomia, professor e atual coordenador do PPGCI/UFF – Mestrado – Programa de Pós-graduação em Ciência da Informação -, professor do Curso de Especialização em Informação Científica e Tecnológica em Saúde, do Ict/Fiocruz, pesquisador do CNPq, desenvolvendo pesquisa em modelização de artigos científicos como ontologias, consultor “ad hoc” da Capes e CNPq, membro do conselho editorial e “referee” de diversos periódicos científicos, membro da Câmara Técnica de Documentos Digitais do Conarq/Arquivo Nacional, atua na área de tecnologia da informação aplicadas ao tratamento de informações, é autor de diversos artigos científicos nessa área; atuou como consultor em projetos como a Biblioteca Digital de Teses e Dissertações – BDTD - do Ibict e no desenvolvimento do servidor SciELO/Open Arquivos.

### *Marília Alvarenga Rocha Mendonça*

Graduada em Biblioteconomia pela Escola de Biblioteconomia da Universidade Federal de Minas Gerais (1996), especialista em Planejamento, Organização e Direção de Arquivos (1985) e Organização e Administração de Bibliotecas Universitárias (1986) pela Universidade Federal Fluminense, Mestre em Administração pela Universidade Federal Fluminense (2001). Atuou como bibliotecária na gerência da Biblioteca da Faculdade de Educação/UFMG (1972/79), na implantação e gerenciamento do Centro de Microfilmagem/UFMG (1979/82), no Núcleo de Documentação/UFF como gerente do Arquivo Geral e da Biblioteca da Faculdade de Farmácia (1982/94). Atualmente é professora assistente do Departamento de Ciência da Informação da Universidade Federal Fluminense. Pesquisadora do Grupo de Pesquisa “Informação, Conhecimento e Tecnologia da Informação”, tendo como linha de pesquisa “gestão de conteúdos digitais e tecnologia da informação”. Participa, desde 2003, de projetos de pesquisa CNPq/UFF, cujos resultados têm sido apresentados em eventos nacionais e internacionais e publicados em Anais e na forma de artigos científicos em periódicos nacionais e internacionais. Possui um livro publicado e três capítulos de livros.