

Relações intra-institucionais na Internet: um estudo exploratório com base em metodologias webométricas

Pamela Barreto Lang

Mestranda em Química Biológica (área de concentração Educação, Gestão e Difusão em Biociências) pela Universidade Federal do Rio de Janeiro e especialista em Comunicação Institucional. Tecnologista em Saúde Pública da Fundação Oswaldo Cruz.

Fábio Castro Gouveia

Doutor em Química Biológica (área de concentração Educação, Gestão e Difusão em Biociências) pela Universidade Federal do Rio de Janeiro. Tecnologista em Saúde Pública da Fundação Oswaldo Cruz.

Jacqueline Leta

Professora adjunta do Instituto de Bioquímica da UFRJ.

Como uma abordagem inicial para um estudo mais aprofundado sobre a imagem e o impacto da Fundação Oswaldo Cruz na Web, este artigo investiga as correlações entre os sites das unidades finalísticas da Instituição localizada no estado do Rio de Janeiro. A metodologia utilizada baseou-se nas análises de co-link e interlink, e na formação de agrupamentos (clustering analysis). Os dados apontaram para quatro agrupamentos distintos e uma Unidade isolada, além de um baixo nível de integração entre os sites estudados.

Palavras-chave: Web; Webometria; Análise de Co-link; Interlink; Site; Fundação Oswaldo Cruz.

Intra-institutional relations on the Web: an exploratory study based on webometric methodologies.

As an initial approach for a more profound study on the image and impact of Oswaldo Cruz Foundation on the Web, this article investigates the correlations between the sites of the units of this institution in Rio de Janeiro. The methodology used was based on co-linking, interlinking and clustering analysis. The data pointed out to the formation of four different groups and one isolated Unit, as well as a low level of integration between the websites studied.

Keywords: *Web; Webometrics; Co-link Analysis; Interlink; Site; Oswaldo Cruz Foundation; Clustering Analysis.*

Recebido em 04.06.2008 Aceito em 13.10.2008

Introdução

As constantes transformações nos cenários político, econômico, cultural e tecnológico representam um grande desafio à comunicação institucional. A transição da era industrial para a era informacional e a ascensão de um novo paradigma tecno-econômico, baseado em informação, inovação e conhecimento, levantam questões delicadas, complexas e multifacetadas para as instituições e seus gestores. A informação e a comunicação adquirem um papel importante como instrumentos e processos poderosos para ampliar as potencialidades estratégicas e a integração das estruturas organizacionais. Esses instrumentos passam a subsidiar e a fazer parte da tomada de decisões, do estabelecimento de parcerias e do relacionamento com clientes/usuários e fornecedores. Dentro dessa lógica, Cardoso (2006) afirma ser necessário que as organizações repensem e aprimorem seus referenciais teóricos e metodológicos tradicionais, a fim de que os processos comunicacionais sejam considerados suportes eficazes para a formulação de suas estratégias institucionais.

As novas tecnologias da informação e, sobretudo, a *Internet*, passam a contribuir para a formação de um novo modelo comunicacional, em que emissor e receptor desempenham papéis ainda mais complexos. Não basta que as instituições estejam na *Internet*; deve-se refletir sobre como elas desejam ser percebidas e de que maneira estão se relacionando com o seu público-alvo.

A partir dessa premissa, está sendo realizada uma pesquisa sobre a imagem institucional na *Web* de uma instituição de ensino e pesquisa de reconhecimento nacional e internacional: a Fundação Oswaldo Cruz (Fiocruz). Sua missão é:

Gerar, absorver e difundir conhecimentos científicos e tecnológicos em saúde pelo desenvolvimento integrado das atividades de pesquisa e desenvolvimento tecnológico, ensino, produção de bens, prestação de serviços de referência e informação, com a finalidade de proporcionar apoio estratégico ao Sistema Único de Saúde (SUS) e contribuir para a melhoria da qualidade de vida da população e para o exercício pleno da cidadania (PLANO..., 2005, p. 2).

Por tratar-se de uma instituição complexa e centenária, a abordagem inicial considerará dados relativos às diversas unidades que a compõe. Este trabalho tem como objetivo, portanto, o mapeamento das inter-relações entre os *sites* das unidades da Fiocruz, utilizando as metodologias de *co-link* e de *interlink*, e análise da possível formação de agrupamentos entre eles. Os dados gerados nesta primeira etapa do trabalho poderão servir de subsídio para uma análise posterior, ainda em andamento, sobre os aspectos da imagem institucional da Fiocruz na *Web*.

2 A Webometria e algumas considerações metodológicas

O termo "webometria" foi cunhado em 1997 por Almind e Ingwersen. Esse novo campo de conhecimento está relacionado às análises quantitativas dentro das áreas de biblioteconomia e da ciência da informação, e surge a partir do desenvolvimento das tecnologias de informação e comunicação. A webometria trata dos aspectos quantitativos, tanto da construção, quanto do uso da *Web*, compreendendo quatro áreas de pesquisa no campo (THELWALL; VAUGHAN; BJÖRNEBORN, 2003): análise de conteúdo das páginas *Web*; análise da estrutura dos *links*; análise do uso da *Web* (por exemplo a análise de arquivos de registro do comportamento de busca e navegação dos usuários); e análise de tecnologias da *Web* (incluindo o desempenho dos mecanismos de busca).

Uma das formas de captura de dados para os estudos webométricos se dá por meio de mecanismos de busca, que conseguem indexar a maior parte das páginas disponíveis na *Web*. Segundo Gulli e Signorin (2005), o Google tem 76% das páginas indexadas, o Yahoo/Altavista tem 69%, o MSN 62%, e o Ask/Teoma 58%, havendo uma sobreposição de 28,85% entre os quatro mecanismos FIG. 1.

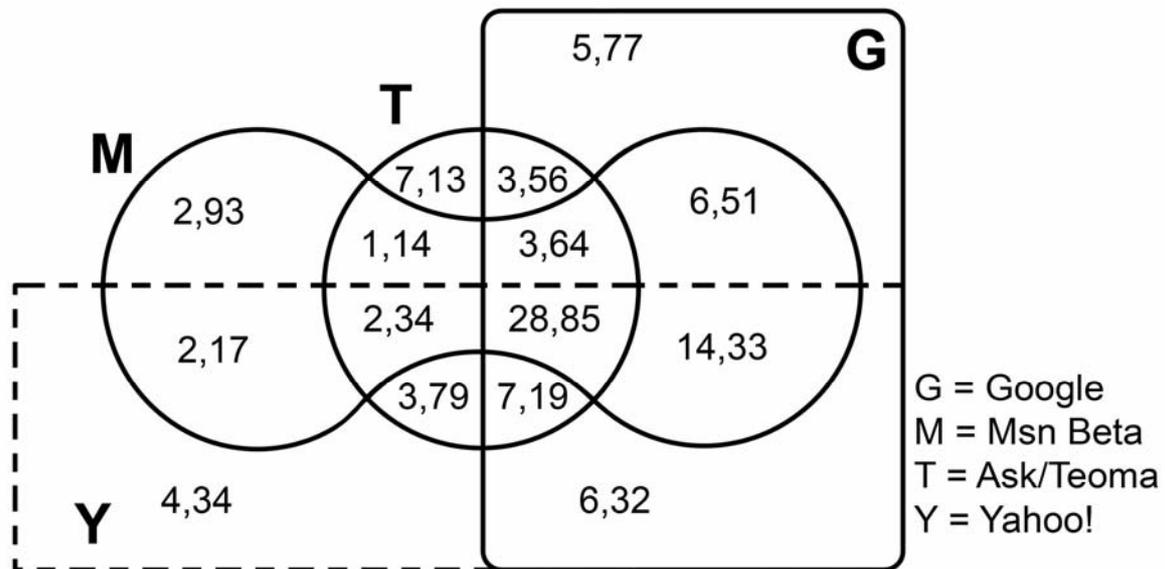


FIGURA 1 - Representação gráfica do percentual de páginas da Web indexadas por mecanismo de busca.

Fonte: GULLI; SIGNORIN, 2005.

A escolha por um determinado mecanismo de busca para a coleta dos dados, geralmente, leva em consideração dois aspectos: o percentual de páginas indexadas e os operadores booleanos disponíveis para a consulta. Os booleanos permitem que sejam realizadas buscas avançadas e sua complexidade varia em cada mecanismo. O *Google*, por exemplo, não é recomendado para buscas combinadas com estruturas complexas. Já o Alta Vista tem sido reconhecido por sua capacidade de busca, possibilitando o uso de cláusulas (*strings*) mais complexas (BJÖRNEBORN; INGWERSEN, 2001).

A indexação das páginas na *Web* sofreu uma mudança radical com o *PageRank*, metodologia criada por Brin e Page (1998), os fundadores do Google, que hierarquiza os resultados de busca de acordo com a rede de *links* voltada para uma determinada página. Ou seja, o método estabelece um ranqueamento com base nas páginas que recebem o maior número de *links* de páginas que também recebem muitos *links*. Dessa forma, quanto maior o número de *links* recebidos por um *site*, maior a sua relevância relativa, o que, de certa maneira, contribui para estimar o valor que lhe é atribuído por outros *sites* (GOUVEIA, 2007). Segundo Vaughan e Thelwall (2005), desde 2002 os mecanismos de busca em geral utilizam alguma implementação similar à descrita por Brin e Page (1998) para o *PageRank*.

Nesse sentido, o *link*, além de se configurar como a unidade central de informação para os estudos webométricos, é considerado um importante indicador do impacto e da posição que determinado *site* ocupa no espaço da rede, já que não apenas representa as relações entre os *sites* e conteúdos disponibilizados, mas também funciona como legitimador desses *sites* para os mecanismos de busca.

Dois métodos vêm sendo utilizados na webometria, especificamente para estimar as relações e similaridades entre *sites*: as análises de

interlink e as de *co-link*. Os *interlinks*, como foram chamados por Thelwall (2002) em estudos sobre universidades britânicas, constituem uma relação de pares, consistindo num método que permite mapear o intercâmbio de *links* entre dois determinados *sites*. Já os *co-links* ocorrem quando uma página (URL) aponta simultaneamente para duas ou mais páginas (VANTI, 2005). Herrero-Solana e Morales-del-Castillo (2004), que chamaram esse método de *co-sitation* – em analogia às co-citações –, consideraram a análise de *co-link* a maneira mais simples de se obter informações sobre as relações entre os *sites*. Vaughan, Kipp e Gao (2007), em um estudo sobre universidades canadenses em que se verificou uma mapa claro de diferenças linguísticas e culturais dentro da sociedade canadense, utilizando a metodologia em questão, afirmaram que os *co-links* podem ser considerados uma medida de similaridade e relação, além de fornecerem dados valiosos para os *sites*.

3 Metodologia

A estrutura organizacional da Fiocruz é hoje composta por 14 Unidades Técnico-Científicas, quatro Unidades Técnico-Administrativas e uma Unidade Técnica de Apoio. Para fins deste estudo, foram selecionados os *sites* de todas as 14 Unidades Técnico-Científicas, por se tratarem de estruturas finalísticas, ou seja, são as unidades que estão diretamente relacionadas ao cumprimento da missão institucional.

Diversos estudos têm encontrado influências geográficas em resultados de análises webométricas. Em um trabalho recente de *co-links* sobre redes de museus e centros de ciência, Gouveia (2007) verificou que a formação de vários agrupamentos estava associada a questões geográficas. Da mesma forma, Thelwall (2002) e Vaughan e Thelwall (2005) também consideraram a proximidade geográfica um fator determinante nas relações entre *sites* de universidades britânicas e universidades canadenses, em seus respectivos estudos.

A fim de isolar tal influência, já que se trata de um mapeamento intra-institucional, não foram incluídos na amostra os *sites* das quatro unidades regionais da Fiocruz, que estão localizadas nos estados da Bahia, Pernambuco, Minas Gerais e Amazônia, permanecendo apenas as dez unidades situadas no estado do Rio de Janeiro. Assim, as análises apresentadas a seguir referem-se às inter-relações de 10 das 14 unidades da Fiocruz, conforme listado no QUADRO 1.

QUADRO1- Nome das dez Unidades Técnico-Científicas da Fiocruz estudadas, com os respectivos endereços da página inicial de seus sites e abreviaturas

Abreviatura	Unidade	Site
BIO	Instituto de Tecnologia em Imunobiológicos	www.bio.fiocruz.br
COC	Casa de Oswaldo Cruz	www.coc.fiocruz.br
Icict	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde	www.cict.fiocruz.br
Ensp	Escola Nacional de Saúde Pública Sérgio Arouca	www.ensp.fiocruz.br
EPSJV	Escola Politécnica de Saúde Pública Joaquim Venâncio	www.epsjv.fiocruz.br
FAR	Instituto de Tecnologia de Fármacos	www.far.fiocruz.br
IFF	Instituto Fernandes Figueira	www.iff.fiocruz.br
INCQS	Instituto Nacional de Controle de Qualidade em Saúde	www.incqs.fiocruz.br
IOC	Instituto Oswaldo Cruz	www.ioc.fiocruz.br
Ipec	Instituto de Pesquisa Clínica Evandro Chagas	www.ipec.fiocruz.br

Fonte: Elaborado pelos autores.

Para a coleta de dados, utilizou-se o mecanismo de busca Alta Vista¹, cuja base de dados é a mesma do Yahoo!² e corresponde a 69% de todas as páginas indexadas na Internet (GULLI; SIGNORIN, 2005).

Na primeira *string*, ou lista das buscas desejadas, foi especificado "link: URLn link: URLn -domain:fiocruz.br" (por exemplo: "link: www.coc.fiocruz.br link: www.ensp.fiocruz.br -domain:fiocruz.br"). Esta busca permitiu coletar todas as listas de páginas que continham *links* para ambas as URLs simultaneamente à exceção daquelas que pertencem ao próprio domínio da Fiocruz (fiocruz.br). Não foi incluído nesta coleta o site do Portal Fiocruz, com o objetivo de analisar a formação de agrupamentos (*clustering analysis*) apenas entre as unidades da Instituição. Essa busca resultou na análise de *co-links* FIG.

2A.

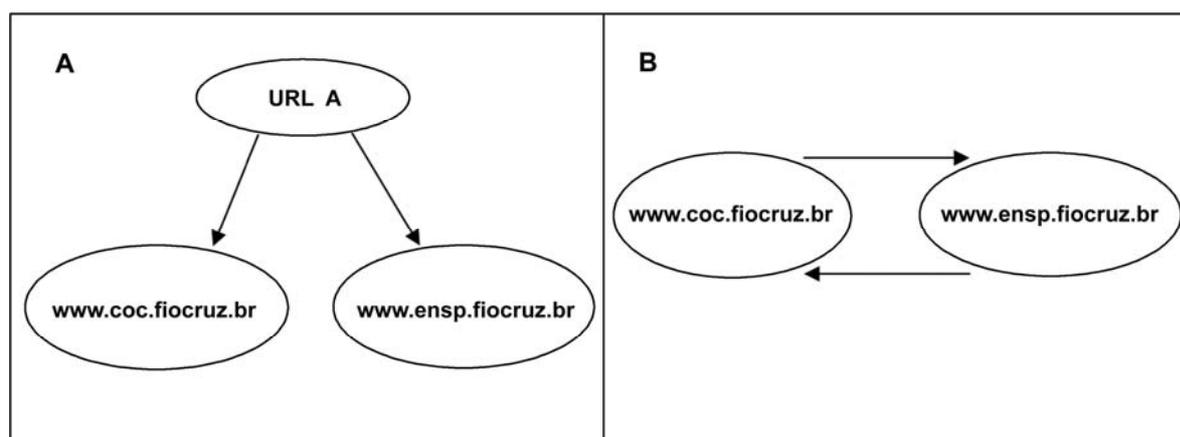


FIGURA 2 - Representação gráfica que exemplifica análises de co-link e de interlink.

Fonte: Elaborado pelos autores.

¹ Disponível em: <www.altavista.com>. Acesso em: 19 jan. 2008.

² Disponível em: <www.yahoo.com>.

Em uma segunda *string* foi especificado "link:URLn host:URLn" (por exemplo: "link:www.coc.fiocruz.br host:www.ensp.fiocruz.br"). Esta busca permitiu coletar as listas de páginas de um determinado *site* (*host*) que contém *links* para a URL inicial da Unidade FIG. 2B.. Nesse caso, considerou-se importante incluir o *site* do Portal Fiocruz, já que a relação entre as unidades e o Portal poderia fornecer dados do ponto de vista da integração e da organização interna dos *links*.

A fim de evitar as possíveis instabilidades na *Web* (BAR-ILAN, 1999) e garantir os melhores resultados possíveis, a coleta de dados aconteceu em um único dia, 19 de janeiro de 2008. Os resultados obtidos foram quantificados e inseridos em duas matrizes.

Uma das questões metodológicas, muito discutida entre os autores do campo da webometria para a análise de *co-link*, é a determinação dos valores usados na diagonal da matriz gerada com os resultados. Zucalla (2006), por exemplo, utilizou a média das colunas para determinar os valores da diagonal. Herrero-Solana e Morales-del-Castillo (2004) e Vaughan e You (2006) consideraram apropriado utilizar os valores correspondentes ao número de páginas do servidor, e Musgrove et al. (2003) optaram por preencher a diagonal com o número zero. No entanto, o presente estudo levou em consideração a abordagem testada por Gouveia e Kurtenbach (2009), que, por meio de método exaustivo, concluiu ser a soma da coluna a melhor opção para lidar com dados de *co-link* numa análise de agrupamento.

As matrizes contendo os resultados obtidos foram então importadas para o programa Statistica 7.0, para o tratamento estatístico dos dados e a análise de agrupamento (*clustering analysis*). O método de Ward (WARD, 1963) foi utilizado para a amalgamação, com a medida de distância de 1 - r de Person.

4 Resultados e discussão

4.1 Análise de *co-links*

A consulta que utilizou a *string* "link:URLn link: URLn - domain:fiocruz.br" gerou uma matriz simétrica de *co-links* entre os dez *sites* da Fiocruz, cujos resultados correspondem ao número de URLs que aponta simultaneamente para ambas as unidades TAB. 1. Os valores da diagonal correspondem à soma das colunas.

TABELA 1- Matriz de *co-links*

	BIO	Icict	COC	Ensp	EPSJV	FAR	IFF	INCQS	IOC	Ipec
BIO	21	1	8	2	1	3	1	1	3	1
Icict	1	15	3	5	2	1	2	0	1	0
COC	8	3	84	55	2	2	3	3	8	0
Ensp	2	5	55	116	3	1	31	2	16	1
EPSJV	1	2	2	3	10	0	0	1	1	0
FAR	3	1	2	1	0	12	1	0	4	0
IFF	1	2	3	31	0	1	48	1	9	0
INCQS	1	0	3	2	1	0	1	10	2	0
IOC	3	1	8	16	1	4	9	2	47	3
Ipec	1	0	0	1	0	0	0	0	3	5

Fonte: Dados da pesquisa.

A partir desses dados é possível estimar a proximidade entre as unidades analisadas. A visualização desta proximidade é melhor percebida no dendograma da FIG. 3, onde as unidades estão ordenadas em grupos (*clusters*). O dendograma, também conhecido por diagrama de árvore, é uma representação gráfica utilizada para a análise de agrupamentos. À esquerda do gráfico estão as unidades analisadas. À medida que se movem no eixo horizontal, que representa uma medida de distância, elas se unem, formando grupos, ou permanecem isoladas, segundo o método de amalgamação utilizado (método de *Ward*), a partir da matriz de distância de $1 - r$ de Person.

No GRÁF. 1, cada nó representa um agrupamento e corresponde a uma distância determinada pelo eixo horizontal. Para comparar os agrupamentos formados, foi necessária uma mesma medida de distância, neste caso de 1,0, a partir da qual podemos observar a formação de quatro grupos, considerando os nós hierárquicos, e uma Unidade isolada.

O primeiro agrupamento contém Biomanguinhos (BIO) e Farmanguinhos (FAR), ambas unidades de produção da Fiocruz, responsáveis pela fabricação de vacinas e medicamentos, respectivamente. A formação deste *cluster* poderia configurar uma relação temática e de conteúdo associada à área finalística delas. E, de fato, ao acessar as páginas que apontam para elas, percebe-se claramente uma motivação temática para o agrupamento.

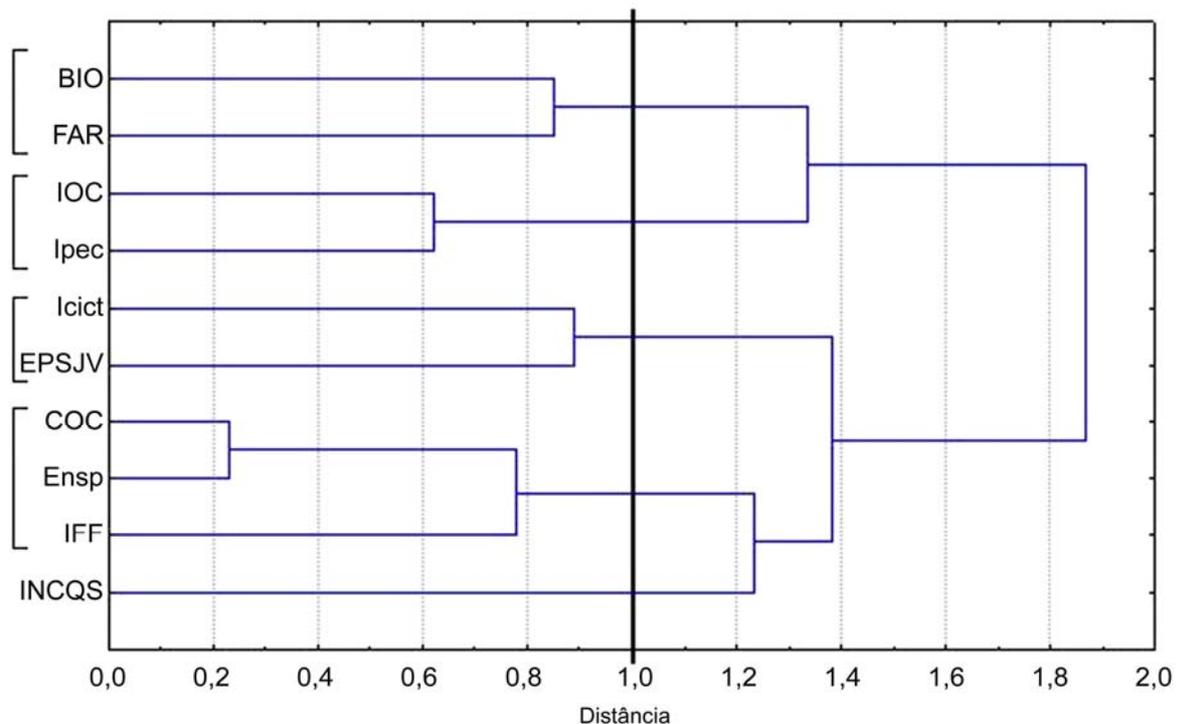


GRÁFICO 1 - Análise de agrupamento a partir da matriz de *co-links*.

Fonte: Dados da pesquisa.

O agrupamento 2 contém o Instituto Oswaldo Cruz (IOC) e o Instituto de Pesquisa Clínica Evandro Chagas (Ipec). A primeira Unidade tem como base de ação seus laboratórios de pesquisa, desenvolvimento tecnológico e inovação, dedicados ao estudo e à geração de produtos e insumos para o combate de diversas doenças; enquanto a segunda é um hospital que presta serviços clínicos e laboratoriais, além de abrigar centros de referência em leishmaniose tegumentar, em micoses e no diagnóstico histopatológico de doenças infecciosas. Embora essas unidades tenham linhas de ação distintas, a formação desse agrupamento pode ser explicada por fatores históricos e políticos. Até 2001 o Ipec era um departamento do IOC e, desde a sua transformação em Unidade Técnico-Científica, tem mantido forte cooperação em quase todas as suas áreas de atuação (por exemplo HIV, hanseníase, dengue, doença de Chagas, leishmaniose e toxoplasmose), inclusive em redes internacionais voltadas para ensaios clínicos de prevenção vertical do HIV, ensaios clínicos de tratamento e de prevenção.

A proximidade temática pode também ser a motivação para o terceiro agrupamento, que é composto pelo Instituto de Comunicação e Informação Científica e Tecnológica (Icict), responsável por atividades de pesquisa, ensino e cooperação técnica nessas áreas, além de investir em novas mídias e tecnologias; e pela Escola Politécnica de Saúde Joaquim Venâncio (EPSJV), que integra atividades de ensino, pesquisa e cooperação técnica, com o objetivo de promover a educação profissional em saúde.

Já o quarto agrupamento reúne três unidades: a Casa de Oswaldo Cruz (COC), a Escola Nacional de Saúde Pública Sérgio Arouca (Ensp) e o Instituto Fernandes Figueira (IFF). Há um forte indicativo de que as duas primeiras estão ligadas também por uma proximidade de temática: a COC caracteriza-se por ser um centro de pesquisa, documentação e informação, dedicado à memória e à história das ciências biomédicas e da saúde pública. Suas principais áreas de atuação contemplam a preservação e a difusão de acervos documentais e museológicos, a pesquisa histórica e o ensino de pós-graduação em história das ciências da saúde. A Ensp está voltada para a prestação de serviços em saúde pública, formação de pessoal especializado em cursos de mestrado e doutorado, e pesquisas nas áreas de promoção da qualidade de vida, prevenção de doenças e ciências sociais aplicadas à saúde.

A presença do IFF neste agrupamento dá-se pela forte relação estabelecida com a Ensp, com 23 páginas que apontam para ambas as unidades simultaneamente. A princípio, esta relação não está voltada para questões relativas à missão institucional, já que o IFF é um hospital de referência na área de aleitamento materno que, embora também realize pesquisas e ofereça cursos de pós-graduação, tem como foco de atuação a assistência médica à criança e à mulher, com destaque para os cuidados de recém-nascidos de alto risco. O acesso às páginas "citantes" nos permitiu verificar que, em quase a sua totalidade, o IFF não está sendo citado como um hospital, mas como instituição de ensino (pós-graduação).

Finalmente, vale notar que o Instituto Nacional de Controle de Qualidade em Saúde (INCQS), responsável pela análise de situações e fatores de risco associados à fabricação e à comercialização de alimentos, medicamentos, cosméticos, produtos biológicos, sangue e seus derivados, aparece isolado das demais unidades, apresentando apenas um *link* com o BIO, três com a COC, dois com a Ensp, um com a EPSJV, um com o IFF e dois com o IOC.

4.2 Análise de *interlinks*

A TAB. 2 apresenta uma matriz assimétrica com os resultados gerados a partir da consulta "link:URLn host:URLn". O eixo vertical representa as unidades que estão cedendo os *links* (ou seja, as unidades que "citam") e o eixo horizontal aquelas que recebem os *links* (ou seja, as unidades que são "citadas"). Os valores da diagonal correspondem à soma das colunas.

TABELA 2 - Matriz de *interlinks*

	BIO	Icict	COC	Ensp	EPSJV	FAR	Fiocruz	IFF	INCQS	IOC	Ipec
BIO	38	0	0	0	0	0	100	0	0	0	0
Icict	0	485	0	2	0	0	474	1	0	1	0
COC	0	0	112	1	0	0	151	0	0	0	0
Ensp	4	15	10	192	7	5	1990	12	1	14	6
EPSJV	0	0	0	0	121	0	30	0	0	0	0
FAR	0	0	0	0	0	39	95	0	0	0	0
Fiocruz	34	470	100	189	114	34	3087	102	34	143	71
IFF	0	0	0	0	0	0	105	115	0	0	0
INCQS	0	0	0	0	0	0	14	0	35	0	0
IOC	0	0	2	0	0	0	92	0	0	160	0
Ipec	0	0	0	0	0	0	36	0	0	2	77

Fonte: Dados da pesquisa.

A baixa proximidade, ou a baixa formação de agrupamentos bem definidos, obtida a partir da análise de *interlinks* entre as unidades da Fiocruz, é melhor observada no GRÁF. 4. Apesar de uma clara proximidade entre o Portal Fiocruz e a Ensp, todas as demais unidades aparecem isoladas, o que demonstra pouca integração entre os seus *sites* (GRÁF. 1). A formação deste agrupamento pode ser explicada pelo fato de tanto a Ensp como o Portal Fiocruz fornecerem um grande número de *links* para as unidades, ou seja, ambas estão dando visibilidade ao conjunto da Fiocruz.

Nota-se ainda que o Portal Fiocruz desempenha um papel central na integração desses *sites*, já que, apesar das unidades, à exceção da Ensp, aparecerem isoladas, todas fornecem *links* para o Portal.

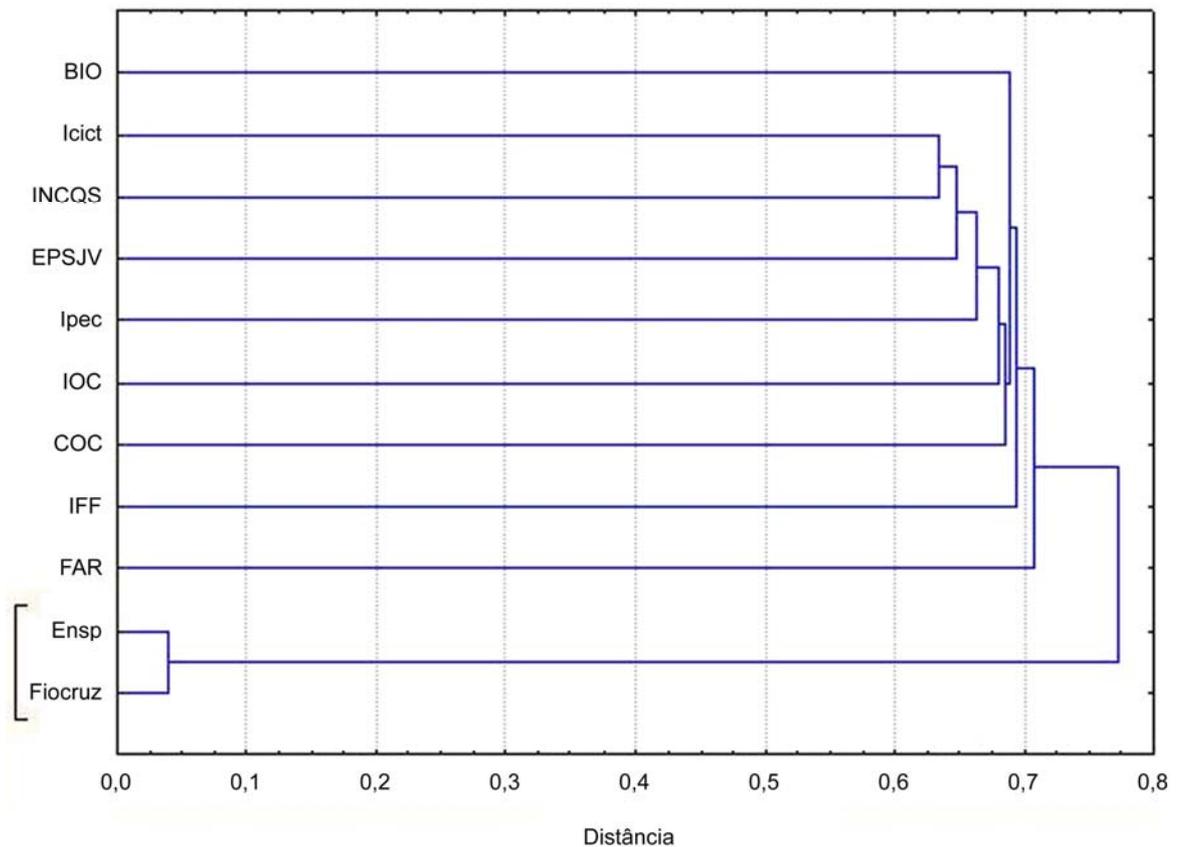


GRÁFICO 2 - Análise de agrupamento a partir da matriz de *interlinks*.

Fonte: Dados da pesquisa.

5 Considerações finais

No que se refere à interpretação dos dados apresentados neste trabalho, há de se considerar algumas das limitações da webometria. Trata-se de um campo de conhecimento recente, ainda em processo de consolidação, e cujas principais metodologias não estão amplamente definidas. Parte da indefinição metodológica está relacionada à instabilidade e à dinâmica de *sites* e *links* na *Web*. Os primeiros estudos realizados a partir de mecanismos de busca mostraram diversas alterações em curtos espaços de tempo, o que tornava os dados bastante instáveis (BAR-ILAN, 1999; ROUSSEAU, 1999; METTROP; NIEUWENHUYSEN, 2001).

No entanto, o fato de todas as coletas terem sido realizadas num único dia minimiza as questões em torno da instabilidade da *Web*, e da validação dos estudos webométricos. Além disso, numa primeira coleta, realizada quatro meses antes, em 19 de setembro de 2007, verificou-se a mesma formação de agrupamentos.

Assim, considerando que encontramos significativa estabilidade de *links* para este recorte institucional, acreditamos que os agrupamentos encontrados a partir dos *co-links* têm influências temáticas. No caso

específico do Ipec e do IOC, a proximidade foi atribuída a questões políticas e históricas. Apesar de difícil identificação, tais motivações poderiam ocorrer também em outros agrupamentos, já que não seria incomum encontrar influências de ordem política, principalmente em grupos intra-institucionais. Herrero-Solana e Morales-del-Castillo (2004), por exemplo, identificaram a formação de um mapa geopolítico ao avaliar a relação entre *sites* de universidades alemãs, norte-americanas e russas.

Há ainda outras questões subjacentes a serem observadas a partir dessas análises, que podem subsidiar um estudo mais detalhado sobre a instituição em questão. O Instituto Fernandes Figueira (IFF) está sendo apontado na *Web* em conjunto com instituições de ensino e pesquisa, e não na categoria hospital, com o Ipec, por exemplo. Talvez o baixo número de *co-links* deste último não tenha permitido um agrupamento temático entre hospitais. Há de se perguntar, no entanto, que público-alvo as unidades desejam atingir ou a ele estar associadas: usuários, estudantes, parceiros institucionais? E, mais, esse fenômeno estaria relacionado a uma estratégia da Unidade ou ocorreria sem o reconhecimento de seus gestores?

No caso da análise de *interlinks*, é possível fazer algumas reflexões a respeito da baixa correlação e integração entre as unidades. Um dos aspectos que poderia exercer alguma influência sobre esses dados seria o surgimento recente do Portal Fiocruz. Apesar do *site* da Instituição existir desde 1995, foi somente em 2005 que ele ganhou uma nova versão, a partir de uma proposta integradora.

O projeto Portal Fiocruz não apenas abarcava um novo *layout* gráfico, mas também pretendia implementar uma política de integração com os *sites* já existentes das unidades. A partir de então, as unidades passaram a incorporar em seus *sites* alguns elementos integradores comuns, por exemplo uma barra superior padrão com *links* para o Ministério da Saúde e a Fiocruz. De fato, embora haja um baixo nível de integração, fica claro o papel central desempenhado pelo Portal Fiocruz, que distribui uma quantidade expressiva de *links*.

A Fiocruz é uma instituição de caráter nacional com interface nas diversas áreas de produção de conhecimento dos sistemas de Ciência e Tecnologia e Saúde. Portanto, seria possível acreditar que os dados encontrados refletem a característica multifacetada da Instituição, na qual cada Unidade tem uma missão e um público-alvo, muitas vezes distintos das demais. Seria recomendável, no entanto, fortalecer as políticas de integração e aumentar a rede de *links* intra-institucionais, a fim de aumentar a visibilidade da Fiocruz no seu conjunto, e legitimar e reconhecer a sua complexidade e a importância de suas áreas finalísticas.

Independentemente das razões por trás dessa conjuntura e das limitações inerentes ao ambiente da pesquisa, vale destacar que estes resultados representam um retrato importante da Fiocruz em suas diversas faces, podendo contribuir para análises futuras e para um plano de reposicionamento estratégico na *Web*.

Referências

- ALMIND, T. C.; INGWERSEN, P. Informetric analyses on the world wide web: methodological approaches to 'Webometrics'. *Journal of Documentation*, v. 53, n. 4, p. 404-426, 1997.
- BAR-ILAN, J. Search engine results over time: a case study on search engine stability, *Cybermetrics*, v. 2/3, n. 1, paper 1, 1999. Disponível em: <<http://www.cindoc.csic.es/cybermetrics/articles/v2i1p1.html>>. Acesso em: 04 out. 2007.
- BJÖRNEBORN, L.; INGWERSEN, P. Perspectives of webmetrics. *Scientometrics*, v. 50, n. 1, p. 65-82, 2001.
- BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, v. 30, n. 7, p. 107-117, 1998.
- CARDOSO, O. Comunicação empresarial versus comunicação organizacional: novos desafios teóricos. *Rev. Adm. Pública*. Rio de Janeiro, v. 40, n. 6, p. 1123-1124, 2006.
- GOUVEIA, F. *Estudos webométricos de associações de museus e centros de ciência*. Rio de Janeiro, Tese (Doutorado em Química Biológica) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2007.
- GOUVEIA, F.; KURTENBACH, E. Mapping the web relations of science centres and museums from Latin America. *Scientometrics*, 2009. [Preprint].
- GULLI, A.; SIGNORINI, A. The Indexable Web is More than 11.5 billion pages. *International World Wide Web Conference*, p. 902-903, 2005. Disponível em: <<http://www.cs.uiowa.edu/~asignori/web-size/size-indexable-web.pdf>>. Acesso em: 23 set. 2007.
- HERRERO-SOLANA, V.; MORALES-DEL-CASTILLO, J. "Geopolitical" maps of the Internet: application of new information representation methods. *Ciência da Informação*, v. 33, n. 3, p. 69-75, 2004.
- METTROP, W.; NIEUWENHUYSEN, P. Internet search engines: fluctuations in document accessibility. *Journal of Documentation*, v. 57, n. 5, p. 623-651, 2001.
- MUSGROVE, P. B. et al. A method for identifying clusters in sets of interlinking Web spaces. *Scientometrics*, v. 58, n. 3, p. 657-672, 2003.
- PLANO Quadrienal da Fundação Oswaldo Cruz - 2005-2008, 2005. Disponível em: <http://www.fiocruz.br/media/plano_quadrienal.pdf>. Acesso em: 14 fev. 2007.
- ROUSSEAU, R. Daily time series of common single word searches in AltaVista and NorthernLight. *Cybermetrics*, v. 2/3, n. 1, artigo 2, 1999. Disponível em: <<http://www.cindoc.csic.es/cybermetrics/articles/v2i1p2.html>>. Acesso em: 04 out. 2007.

THELWALL, M. Evidence for the existence of geographic trends in university web site interlinking. *Journal of Documentation*, v. 58, n. 5, p. 563–574, 2002.

THELWALL, M., VAUGHAN, L.; BJÖRNEBORN, L. Webometrics. *Annual Review of Information Science and Technology (ARIST)*, v. 39, 2003. [Preprint].

VANTI, N. Os links e os estudos webométricos. *Ciência da Informação*, Brasília, v. 34, n. 1, p.78-88, jan./abr. 2005.

VAUGHAN, L., THELWALL, M. A modeling approach to uncover hyperlink patterns: the case of Canadian universities. *Information Processing and Management*, v. 41, p. 347–359, 2005.

VAUGHAN, L., YOU, J. Comparing business competition positions based on Web co-link data: the global market vs. the Chinese market. *Scientometrics*, v. 68, n.3, p. 611-628, 2006.

VAUGHAN, L.; KIPP, M.; GAO, Y. Why are Websites co-linked? The case of Canadian universities, *Scientometrics*, v. 72, n. 1, p. 81-92, 2007.

WARD, J. H. Hierarchical Grouping to optimize an objective function. *Journal of American Statistical Association*, v. 58, n. 301, p. 236-244, 1963.

ZUCCALA, A. Author cocitation analysis is to intellectual structure as Web co-link Analysis is to...? *Journal of the American Society for Information Science and Technology*, v. 57, n.11, p. 1487-1502, 2006.