

# ACERVOS DIGITAIS DE JORNAIS: CONSIDERAÇÕES A RESPEITO DA RECUPERAÇÃO DA INFORMAÇÃO SOBRE FUTEBOL

**Luana Carla de Moura dos Santos**

Mestre em Ciência da Informação pela Universidade Federal de Santa Catarina (UFSC), graduada em Biblioteconomia pela Universidade do Estado de Santa Catarina.

E-mail: luana-moura@hotmail.com

**Marisa Bräscher Basílio Medeiros**

Professora do Departamento de Ciência da Informação da Universidade Federal de Santa Catarina.

E-mail: marisa.brascher@gmail.com

**Resumo:** Apresenta o jornal como importante fonte de informação para pesquisa histórica de uma sociedade. Aborda os acervos digitais, ambientes que funcionam como arquivos *on-line* de jornais como possibilidade para acessar de forma mais facilitada esses documentos. Discute como a recuperação da informação deve ser aperfeiçoada para possibilitar que essa informação seja acessada de forma facilitada e dependa menos do letramento do usuário. Aponta que essa recuperação deve considerar variações linguísticas e socioculturais, além da especificidade de cada domínio do conhecimento. Para tal, expõe a importância do jornal para o futebol, as lacunas existentes na organização dessas informações nesse domínio do conhecimento e apresenta a ontologia como um sistema de organização do conhecimento eficiente para favorecer os acervos digitais de jornais. Ao direcionar as ontologias para um domínio, nesse caso o futebol, mais específica ela se torna, fazendo com que as ambiguidades desse contexto sejam tratadas conforme as suas reais finalidades.

**Palavras-chave:** Acervo digital de jornal. Futebol. Organização e Recuperação da Informação.



# 1 INTRODUÇÃO

O surgimento da prensa de Gutenberg viabilizou a multiplicação dos documentos impressos, a expansão da Internet estimulou as publicações em meio digital. Nessa transição do impresso para o digital, suscitada pelo avanço significativo das Tecnologias da Informação e Comunicação (TIC), foi evidenciada a necessidade de disponibilizar digitalmente os documentos que foram produzidos em suporte de papel, em décadas e séculos passados.

Essas iniciativas surgiram primeiramente em ambientes internos de bibliotecas, instituições e organizações. As informações eram armazenadas em bancos de dados privados, utilizados em princípio para suprir as carências informacionais de uma reduzida comunidade. Subsequentemente, com o intensivo uso da Internet foi evidenciada a necessidade da disponibilização desses bancos de dados *on-line*, possibilitando assim que um maior número de pessoas tivesse acesso a essas informações.

No final do século XX e início deste século, jornais brasileiros que mantinham internamente seus bancos de dados, disponibilizaram seus acervos digitais na Internet. O jornal em ambiente digital oferece vantagens quando comparado à versão convencional. O acesso não é limitado por um único leitor por exemplar, conforme a versão em papel. Permitindo, dessa forma, que vários leitores possam fazer uso simultaneamente de um mesmo documento. A guarda e preservação são facilitadas pelos bancos de dados dos próprios jornais e a pesquisa é simplificada pelo auxílio de recursos que propiciam encontrar termos nos textos, por meio de uma navegação intuitiva que permite a busca por palavras-chave em seções do jornal (esportivo, econômico, policial) e períodos de tempo (dia, mês, ano) (SANTOS; VIERA, 2015).

Essas transformações estão em consonância com a convergência midiática, termo cunhado por Henry Jenkins para designar a tendência de adaptação dos meios de comunicação à

Internet. De acordo com Jenkins (2008, p. 41), “palavras impressas não eliminaram as palavras faladas. O cinema não eliminou o teatro” e, neste cenário, meios de comunicação são adaptados devido à necessidade de introdução de novas tecnologias.

No processo de convergência de mídias, além do fluxo dos conteúdos serem alterados, o comportamento do usuário também é migratório, “a convergência envolve uma transformação tanto na forma de produzir quanto na forma de consumir os meios de comunicação” (JENKINS, 2008, p.44). Portanto, para acesso aos acervos digitais de jornais, além de oferecer os conteúdos da antiga mídia, é preciso habilitar os recursos fundamentais praticados na nova mídia. Seus padrões de acesso, organização e, principalmente, recuperação da informação (RI) precisam ser adaptados.

Outra questão que se coloca do ponto de vista da RI em acervos digitais que reúnem documentos de diferentes épocas e culturas é a variação linguística que ocorre em função da forma diacrônica “(dizer já dito em momentos diferentes)” e sincrônica “(dizer atual e simultâneo)” (MOURA, 2009, p. 64), assim como aqueles referentes aos contextos de uso, marcados muitas vezes por aspectos geográficos e socioculturais.

Essas problemáticas repercutem de forma intensa em alguns domínios do conhecimento, como o futebol, por exemplo, que é uma atividade com mais de 120 anos no Brasil. Nessa passagem, o jornal foi uma mídia fundamental, em que contextualiza e narra acontecimentos do esporte, apresentando assim, transições na gramática, além de, pelo fato de ser um esporte praticado e acompanhado por um público grande, diverso, e geograficamente disseminado nas regiões brasileiras, apresenta terminologias distintas que designam um mesmo conceito. Assim, esse artigo busca abordar a representatividade dos jornais para o futebol, e como a recuperação da informação, considerando aspectos da pluralidade e transição da gramática, devem ser levados em conta em ambientes similares aos acervos digitais de jornais.

## 2 PROCEDIMENTOS METODOLÓGICOS

A pesquisa busca, com base na literatura, apresentar pontos relevantes sobre a importância dos jornais e a organização de acervos digitais que contemplam essa mídia. Isso é feito a partir da perspectiva de um domínio do conhecimento significativo e representativo para o país, o futebol. Assim, a pesquisa em questão é bibliográfica. De acordo com Lakatos e Marconi (2012, p.57), “a pesquisa bibliográfica, ou de fontes secundárias, abrange toda bibliografia já tornada pública em relação ao tema de estudo, desde publicações avulsas, boletins, jornais, revistas, livros, pesquisas, monografias, teses, material cartográfico, etc”.

Quanto à forma de abordagem a pesquisa é qualitativa. Para Triviños (2010, p. 125), a pesquisa qualitativa possibilita ao pesquisador compreender o “contexto do fenômeno social que se estuda, privilegia a prática e o propósito transformador do conhecimento que se adquire da realidade que se procura desvendar em seus aspectos essenciais e acidentais”.

O trabalho também pode ser considerado como de revisão, de acordo com Alvarenga e Silva (2010), esta tipologia representa tentativas de apresentar a fase corrente de uma pesquisa de determinado recorte particular de um campo científico, possibilitando informar os mais recentes avanços, lacunas e ideias sobre o assunto.

Para apoiar as discussões e apresentar exemplos concretos foi utilizado como base o Acervo digital Estadão<sup>1</sup>, que funciona como um arquivo *on-line* de todas as edições impressas do Jornal “O Estado de S. Paulo”, desde 1875. Também fazem parte as edições mais recentes que nasceram digitais e que ficam disponíveis após 30 dias da data de publicação. O Acervo O

---

<sup>1</sup> Acervo Estadão: Disponível em: <http://acervo.estadao.com.br/>. Acesso em: 21 out. 2016.

Globo<sup>2</sup> também foi empregado, pois permite o acesso à versão digital de todas as páginas e matérias do Jornal O GLOBO desde sua primeira edição, em 29 de julho de 1925.

Desta forma, para contextualizar a problemática, são levantados tópicos sobre patrimônio digital, a transição do jornal impresso ao digital, e as tratativas que envolvem informação, jornal e futebol, e a importância de manter ambientes como os acervos digitais organizados, a fim de possibilitar que a pesquisa nessa mídia, e nesse domínio do conhecimento seja facilitada e viável.

### 3 PATRIMÔNIO DIGITAL

Para compreensão inicial de patrimônio, que está diretamente ligado ao conceito de herança, precisamos definir os conceitos de memória, registro, testemunho e informação, que são componentes que interferem significativamente nas evoluções históricas, e são elementos que integram o Patrimônio digital.

O sentido original de memória refere-se à predisposição humana em guardar no cérebro as impressões das experiências vividas (LOPES, 1998). De acordo com Chapouthier (2006, p. 9) “memória é a capacidade que certos seres vivos têm de armazenar, no sistema nervoso, dados ou informações sobre o meio que os cerca, para assim modificar o próprio comportamento”.

No entanto, essa acepção de memória, em relação ao indivíduo, concebida de forma natural e particular, é imediata e vulnerável. Para Lévy (2000, p.177), essa afirmação pode ser entendida pelo trecho “[...] quando um ancião morre, é uma biblioteca que se queima”, pois com ele carrega todas as recordações armazenadas somente em sua memória.

---

<sup>2</sup> Acervo O Globo: Disponível em: <http://acervo.oglobo.globo.com/>. Acesso em: 22 out. 2016.

Quando recorremos ao entendimento de “memória social” (LE GOFF, 1990), é possível estabelecer o raciocínio sobre a memória associada ao documento, que pode ser aplicada ao fundamento da memória sendo registrada, ou pautada na comunicação e na troca de informações (RUEDA; FREITAS; VALLS, 2011), que nesse caso, não necessariamente será registrada, mas permanecerá sendo disseminada devido aos testemunhos vivos que os perpetuam.

Para Le Goff (1990), a ação de perpetuação da história de uma sociedade pode ser compreendida como características de monumento e documento. O primeiro, o monumento, advém do entendimento que a preservação da memória, voluntária ou involuntária da sociedade, é um legado coletivo, em menor parcela de testemunhos escritos. E o segundo, o documento, ganha caráter de ‘prova’, sendo considerado o fundamento do fato histórico.

A memória é um conjunto de todos esses movimentos, da memória como testemunho e registro, e a informação será a extração desse conteúdo, que de maneira geral sobrevive devido àqueles que se dedicam a ciência e desenvolvem ações tendo como a motivação a preservação da história de uma sociedade para tempos futuros, constituindo os patrimônios históricos.

Para Dodebei (2006), patrimônio pode ser considerado como um atributo simbólico de um bem que se deseja preservar com indício de memória. Le Goff (1990) corrobora ao afirmar que desde a antiguidade, a ciência histórica reúne documentos escritos e faz deles testemunhos. com a transformação da sociedade, as formas de registramos as informações foram alteradas. Inicialmente tinha-se a forma oral, sucedidas pelas inscrições rupestres na pré-história, feitas geralmente em pedras, madeiras ou cavernas. Após ocorreu a revolução tecnológica em 1450, e o surgimento dos tipos móveis e a prensa de Gutenberg, aí veio à chamada “explosão documental” e atualmente ainda estamos em transição para o ciberespaço (RUEDA; FREITAS; VALLS, 2011).

Nessa transição, podemos considerar o jornal como indicio de memória, na forma testemunhada, por aqueles que compartilham suas impressões sobre assuntos pertinentes em uma determinada época, como na forma registrada, por meio da escrita, e em decorrência desses traços é evidenciada a importância do jornal para o fornecimento de informações sobre a história de uma sociedade.

Le Goff (2003, p. 25) afirma que “[...] o passado é uma construção e uma reinterpretação constante, e tem um futuro que é parte integrante e significativa da história”. Precisamos contemplar uma nova erudição, que responda de forma simultânea às exigências tecnológicas e a influência que elas detêm sobre a memória coletiva.

Assim, se as formas de registro da memória evoluíram, conseqüentemente, as formas de acessá-las foram modificadas. Dessa maneira, ainda em passagem, os documentos que dispõem da memória registrada, que antes eram exclusivos de poucos pesquisadores, estão sendo migrados para ambientes digitais e assim sendo acessíveis a um número maior de pessoas. Esse cenário é possível com a introdução das TIC. Nesse contexto, fontes históricas continuam em desenvolvimento, acompanham as tecnologias e se tornam patrimônios digitais.

Por meio da Declaração Universal sobre a Diversidade Cultural, do Programa Memória do Mundo, a UNESCO (2003) lançou a Carta sobre a preservação do Patrimônio Digital. A carta manifesta a preocupação dos países sobre a salvaguarda da memória mundial registrada em documentos digitais. As ações tratadas estão iniciativas que devem ser consideradas que vão desde a criação destes documentos, até às políticas de acesso, as quais devem ser constantemente revistas.

No acesso, é preciso aplicar as práticas de indexação, arquivamento além da reprodução dos documentos em diferentes mídias. Porém, convergir o meio impresso para o ambiente digital, com a inclusão dos documentos “nascidos digitais”, a fim de tornar esses ambientes receptivos para usuários ubíquos, não é

uma tarefa simples. Por esta razão, a intervenção contínua, com a atualização de tecnologias, vocabulários e métodos de organização da informação se faz necessária.

No contexto digital, o patrimônio, de acordo com Dodebei (2006), pode ser estudado sobre dois enfoques: o primeiro considera o patrimônio digital como um processo de constante produção, seja dos objetos digitalizados, ou dos objetos já que já nasceram digitais (*born-digital heritage*). O segundo refere-se ao patrimônio como produto de uma escolha e, nesse enfoque, um conjunto isolado de objetos é escolhido para representar, de forma simbólica, as ações sociais.

Os acervos digitais de jornais estão em conformidade com o primeiro enfoque, pois disponibilizam em seus sítios tanto edições de jornais que foram criados e distribuídos em papel, quanto edições que nasceram digitalmente. Ambos os formatos se integram e formam uma linha do tempo de informações históricas que possuem data de início, assinalada pelos primeiros exemplares dos jornais, mas não possuem data final, pois a história continua e a reprodução dos jornais segue concomitante.

O segundo enfoque é colocado em evidência e se relaciona com a salvaguarda da integridade dos acervos. No sentido digital, além da preservação, as formas de representação que garantem o acesso necessitam ser contempladas.

Para tal, Dodebei (2011) enumera outras cinco questões que devem ser levadas em consideração na passagem da sociedade analógica para a digital:

1. Salvaguardar ainda a produção intelectual armazenada nas memórias documentárias
2. Fazer os patrimônios digitais conversarem entre si;
3. Considerar a leitura hipertextual dos recursos organizados em páginas, sites e portais na web;
4. Indexar e recuperar os recursos informacionais com auxílio de programas lógicos e manter arquivos reproduzíveis;
5. Proteger ou disseminar (como forma de preservação dinâmica) as memórias que circulam na web.

O cenário abordado neste artigo está diretamente ligado à quarta questão, voltada à construção de modelos que incluam a lógica na sua representação, com o objetivo da recuperação da informação. Para Arellano (2004, p.18) “o dado e a mídia que suportam a informação devem possuir um nível de funcionalidade representacional que permita a sua reprodução a qualquer momento que a instituição mantenedora precisar recuperar o dado”.

No que se refere aos acervos digitais, de acordo com Dodebei (2006, p.9), a virtualização e a digitalização são demandas práticas e cotidianas no processo de transferência da massa documental física para ambientes digitais. Cavalcante (2007) afirma que disponibilizar uma parte do patrimônio histórico-cultural digitalmente, é uma forma de coletivizar acervos que antes eram destinados apenas a preservação, e não ao acesso, devido as fragilidades provocadas pela degradação física provocadas por séculos de existência.

Contudo, os novos suportes carecem de intervenção para o desenvolvimento de mecanismos que propiciem um acesso facilitado a essas mídias e, nesse sentido, explorar medidas que visem o uso contínuo de informações, para que elas não se tornem inacessíveis, também é fundamental.

Os acervos digitais se constituem de documentos históricos, por isso apresentam particularidades em seu vocabulário, que devem ser levadas em consideração quando organizados. Quando se trata de um contexto dinâmico é preciso aplicar métodos de representação que permitam contemplar essas particularidades, como as variações linguísticas com aspectos diacrônicos e sincrônicos, a fim de que essas informações sejam relacionadas e oportunizem uma recuperação da informação mais integrada, unindo termos que possuem o mesmo significado, mas formas de grafia diferenciada, devido às mudanças de ortografia ou aos estrangeirismos encontrados nos acervos de jornais.

### **3.1 Patrimônio: o jornal do impresso ao digital**

O jornal, após o livro, é o veículo impresso mais longevo e, assim como o livro, sobreviveu a todas as mudanças tecnológicas e sociais transcorridas na história. Conforme Dines (2001, p.26) o jornal resistiu ao “obscurantismo político, à revolução industrial, às invenções para vencer distâncias, às guerras, às migrações, à recensão econômica, à TV” e todas essas ocorrências foram absorvidas e estão registradas nas páginas dos jornais diários.

De acordo com informações retiradas da cronologia da Associação Nacional dos Jornais (ANJ, 2008)<sup>3</sup>, a história do jornal no mundo inicia-se em 59 a.C., e sua invenção está atribuída ao Imperador Romano, Júlio César. Na ocasião, o Imperador ordenou que os principais acontecimentos políticos do dia fossem publicados e divulgados para seu império em um documento chamado de Acta Diurna.

Na Idade Média, o jornal teve o seu grande impacto tecnológico viabilizado pela prensa de Gutenberg, no ano de 1447, o que acarretou numa maior disseminação e audiência. A partir desse momento, surge a possibilidade da produção em massa, o que determinou maior acesso pela população, de forma mais rápida e ampla.

Após o surgimento da prensa de Gutenberg, importantes países tiveram a publicação dos seus primeiros jornais. No Brasil, o impresso nasceu no início do século XIX (1808) com finalidades políticas, diferentemente dos interesses de fortalecimento da classe mercantil, demonstrados no cenário Europeu, no início do século XVII, e nas Américas espanhola e inglesa, ao final do mesmo século (CALDAS, 2004, p.42).

No entanto, no ano de 1844, outra importante contribuição para o jornal e para a comunicação foi inventada, o telégrafo.

---

<sup>3</sup> Linha do tempo com a cronologia da história do jornal no Mundo. Disponível em: <http://www.anj.org.br/cronologia>. Acesso em: 22 set 2016.

O aparelho revolucionou a forma de se comunicar à distância, viabilizando a transmissão de mensagens de um ponto a outro. Conseqüentemente, a imprensa teve uma maior abrangência e a distribuição de notícias à sociedade foi facilitada.

Os jornais brasileiros dispunham, além do frescor dos acontecimentos diários, o espírito crítico da corrente literária. Famosos autores como Machado de Assis, Inglês de Souza, Aluísio de Azevedo, Júlio Ribeiro e Raul Pompéia descreveram suas visões sobre os aspectos sócio econômicos, culturais e políticos vivenciados na época (FERRARI, 2006, p.26). Tal era a importância atribuída ao jornal, que em um artigo de Machado de Assis, publicado originalmente no Correio Mercantil, no ano de 1859, repercutia as incertezas do autor em relação a sobrevivência do livro perante ao triunfo do jornal - “O jornal matará o livro? ” - foi uma das perguntas levantadas pelo autor. O artigo também colocava em questão a revolução contida na imprensa e que o jornal podia ser considerado um “banco intelectual”:

O jornal, operando uma lenta revolução no globo, desenvolve esta indústria monetária, que é a confiança, a riqueza e os melhoramentos. O crédito tem também a sua parte no jornalismo, onde se discutem todas as questões, todos os problemas da época, debaixo da ação da idéia sempre nova, sempre palpitante. O desenvolvimento do crédito quer o desenvolvimento do jornalismo, porque o jornalismo não é senão um grande banco intelectual, grande monetização da idéia, como diz um escritor moderno (ASSIS, 1859).

Ambos, livro e jornal, suportaram as transformações ocorridas na evolução da sociedade, porém, necessitaram acompanhar as mudanças, principalmente as advindas da tecnologia. Em 1880 as primeiras fotografias foram associadas as páginas dos jornais, até então a mídia era somente constituída de textos.

No início do século XX, o jornal ainda era o principal veículo de comunicação da sociedade, fato que culminou o título de anos dourados à mídia (ANJ, 2008), mas, nos anos 1920, surge o rádio, e os jornais tiveram que se remodelar para compreender o seu real papel na sociedade.

Com a popularização do rádio, ocasionada pela sua facilidade de consumo, comparada ao jornal, que precisava de uma atenção maior do leitor, foi repercutido que os editoriais da mídia impressa seriam extintos. No entanto, os editores se adaptaram às novas tecnologias e apresentaram um produto com mais textos e imagens, a fim de inovar os formatos já existentes para continuar atraindo o público (ANJ, 2008)<sup>4</sup>.

Logo após o surgimento do rádio, ainda nos anos 1920, surge a televisão, potente mídia de massa, que além de contar com o sonoro do rádio, ainda emitia imagens em ação, o que configurava um aparelho de forte entretenimento para sociedade. Neste período, o jornal tinha se readaptado ao surgimento do rádio, porém, a eclosão da televisão fez com que novas estratégias fossem adotadas para que a mídia impressa não se tornasse obsoleta. A adoção de textos curtos, coloridos e diretos, foram aplicados como alternativa para a concorrência com a televisão.

Na década de 1960 surgiu a Internet. Uma rede de computadores chamada primeiramente de Arpanet, desenvolvida pelo Departamento de defesa dos Estados Unidos para fins militares. (CASTELLS, 2003, p.13). Posteriormente, no final da década de 1980 e início da década de 1990, surge a *World Wide Web*, desenvolvida por Tim Berners- Lee. Com o surgimento do “www” a Internet se difundiu globalmente de forma mais acessível e organizada, o que permitiu ser considerada também como um sistema de comunicação. (CASTELLS, 2003, p.8;18)

A Internet, também intitulada “rede” passou por uma imensa transformação entre os anos 60 e 90 do século passado

---

<sup>4</sup> Jornais: Breve História. Disponível em: <<http://www.anj.org.br/jornais-breve-historia/page-2?showall=>>>. Acesso em: 08 set 2016.

(VAZ, 2001), no entanto, o “boom” da Internet ocorreu nos anos 1990, com o progresso da Web.

Com a eclosão da Internet, os jornais precisaram adentrar nesse contexto digital, assumindo assim, também, outras características, mas não perdendo a sua identidade social e cultural, de informar. De acordo com Tellaroli e Albino (2007, p.13), a cada transformação tecnológica, a produção na imprensa precisava se transformar também, para seguir a evolução que a sociedade exigia:

Nota-se que a imprensa transforma-se a cada inovação tecnológica; muda a maneira de produção de conteúdo e altera a forma de recepção. Exemplos: o veículo impresso caracteriza-se por exigir do público o conhecimento da escrita; o rádio atinge um número maior de pessoas por usar apenas o som (audição); a televisão atrai um número também elevado de pessoas pelo apelo visual e auditivo; já o veículo on-line agrega em si todas as características dos outros veículos e gera uma mudança de paradigma da tecnologia da informação.

Com a introdução dos jornais na web, um dos seus traços principais é perdido - o papel. No ambiente digital é preciso salientar a diferença de duas perspectivas do jornal: O *on-line* e o digital. A dimensão *on-line*, de acordo com Palacios (2003) apresenta cinco características: Multimídia/ convergência, Interatividade, Hipertextualidade, Personalização e Memória. Nessa perspectiva, ainda é possível incluir a característica da instantaneidade.

A multimídia e convergência estão relacionadas à conversão de imagem, som, texto das notícias jornalísticas para diferentes sistemas. A interatividade está voltada à possibilidade de o leitor participar de forma colaborativa no processo de criação das notícias ou na inclusão de críticas em notícias já publicadas.

A hipertextualidade traz ao leitor a possibilidade de percorrer o texto de uma forma não linear, viabilizando, por meio dos hiperlinks, que o leitor siga caminhos diferentes e chegue a informações distintas. A personalização contempla as necessidades de cada leitor, oportunizando assim que os jornais se adaptem, de acordo com as individualidades de cada leitor. A memória surge com a compreensão de que tudo o que for criado digitalmente, será preservado para acesso a longo prazo e será possível sua recuperação. Por fim, a instantaneidade permite a atualização dos conteúdos disponibilizados na web, conforme o ritmo dos acontecimentos, isso possibilita o acompanhamento contínuo dos assuntos de maior interesse pelos leitores.

Na dimensão digital, termo que adotamos aqui para discorrer sobre o jornal disponibilizado nos acervos digitais, as características são as mesmas da versão impressa, com publicação de nova edição feita a cada 24 horas, sendo essa, disponível primeiramente apenas para assinantes. Após um período estipulado pela gestão dos acervos, geralmente 30 dias, a edição do jornal é disponibilizada no banco de dados composto pelas versões mais antigas.

Os acervos digitais de jornais fundamentados na dimensão digital, são compostos tanto por jornais digitalizados e jornais que já nasceram digital, os chamados “patrimônio nascido digital” (*born-digital heritage*), abordados nessa pesquisa como “nascidos digitais”, juntos os dois formam os objetos de estudo desta pesquisa. As características abordadas na dimensão *on-line* não farão parte das pesquisas desse estudo, exceto a característica referente à memória, que também é aplicada à dimensão digital.

Um dos traços mais importantes que possibilitaram o surgimento dos acervos digitais de jornais, foi a digitalização. De acordo com Schiwingel e Zanotti (2010, p. 86) a prática da digitalização no jornalismo ocasionou uma transformação profunda:

A referência à digitalização como fenômeno específico da expressão tecnológica sobre o jornalismo procura demarcar o início de uma transformação profunda na atividade jornalística em que os registros captados ou gerados na organização jornalística são configurados em um formato digital, o que garante uma conversibilidade e trânsito dos conteúdos em diferentes plataformas computacionais.

A possibilidade de transferir tudo o que foi criado em papel, anteriormente à propagação da Internet, para as telas de computador foi uma conquista para os acervos digitais. No caso dos jornais, foi possível deixar acessíveis séculos de história. Para Montenegro e Silva (2005), um dos benefícios que a evolução das TIC permitiu foi a possibilidade de utilização de informações e imagens dos documentos originais por meio da digitalização, além da facilidade de armazenar, recuperar e distribuir em meio digital essas informações, sem que os documentos originais estejam em constante manuseio, procedimento que prolonga a sua existência e auxilia a preservação.

Os acervos digitais são compostos em sua maioria por documentos digitalizados, pois a produção de jornais em ambiente digital teve início em meados da década de 90 do século XX, e os jornais já vinham sendo publicados no Brasil desde o século XIX. Com a junção dos documentos nascidos digitais, os acervos de jornais compõem uma base de dados de contínuo crescimento, fortalecendo a importância dessa fonte para pesquisa histórica.

Dessa forma, é importante assegurar que os procedimentos e recursos de organização da informação levem em consideração as características intrínsecas dos acervos digitais, permitindo a relação semântica de forma mais abrangente que a organização tradicional, por palavras-chave. Quando essa organização levanta conceitos relevantes para determinado domínio, pode contribuir para que a RI seja mais precisa.

O domínio do futebol, explorado nessa pesquisa, apresenta muitas das particularidades expostas nos acervos digitais, como as mudanças ortográficas, e também traz consigo as características derivadas do seu vocabulário, com aspectos da sinonímia, polissemia, homografia, palavras derivadas do inglês que foram modificadas para o português no passar dos anos, expressões que necessitam de inferência para serem compreendidas, além dos neologismos, que são responsáveis por atribuir novos sentidos a palavras que já existem na língua. Portanto, o domínio do futebol apresenta um cenário diferente nos acervos digitais, do que quando exposto em outros Sistemas de Recuperação da Informação (SRI), o que evidencia a necessidade de representação de diferentes tipos de relações entre os conceitos do domínio, por vezes não contempladas em vocabulários controlados.

### ***3.2 Informação, Futebol e Jornal***

Neste tópico não abordaremos como se constrói a informação esportiva, quais são os seus usuários e fontes de coleta. Mas sim, como se constituem as informações derivadas do futebol, contextualizadas nas notícias esportivas provenientes de jornais impressos, a fim de realçar a importância dessas informações para o campo de pesquisa.

O Brasil é conhecido popularmente como o “País do futebol”. Seus consumidores, chamados de torcedores, acompanham diariamente o esporte, discutem, comentam, lotam estádios e tornam-se clientes de suas marcas.

De acordo com Guterman (2009) o futebol foi introduzido no Brasil por Charles Miller há mais de 120 anos, um esporte de origem inglesa, praticado inicialmente por homens da elite e brancos. Os primeiros clubes a surgirem no país, praticavam um esporte amador, sem regras bem definidas, estilista, excludente e racista (JORGE, 2013).

Conforme a cronologia de Máximo (1999, p.2), o período que contempla o nascimento dos primeiros clubes no Brasil, vai de 1896 a 1905:

Em 1896, o São Paulo Athletic Club, fundado oito anos antes, seria o primeiro a aderir ao novo esporte, logo seguido do Sport Club Germania (1889), de Mackenzie Athletic Association (1898), Sport Club Internacional (1898), Clube Atlético Paulistano (1900), já com nome aporuguesado. Em Campinas, fundou-se a Associação Atlética Ponte Preta (1900). No Rio de Cox, o Fluminense Foot-ball club (1902), o Rio Foot-ball Club (1902), o Botafogo Foot-ball Club, o América Foot-ball Club, o Bangu Athletic Club (os três últimos em 1904). Flamengo e Vasco da Gama já existiam desde o fim do século, ambos dedicando-se ao remo: o primeiro, só criaria seu departamento de futebol em 1911; o segundo, em 1923. Em Porto Alegre, foi fundado o Esporte Clube Rio Grande (1900); em Minas, o Sport Club Belo Horizonte (1904); em Recife, o Club Náutico Capeberibe (1901); em Salvador, o Vitória Foot-ball Club (1905).

Os clubes mais antigos do país, ainda em atividade, são o Sport Club Rio Grande, do Rio Grande do Sul e a Associação Atlética Ponte Preta, ambos fundados em 1900. Em 1933, o futebol brasileiro passou a ser profissional. A regulamentação foi assinada pela legislação social e trabalhista do Governo de Getúlio Vargas, e assim foi criada a profissão de jogador de futebol (CARRAVETTA, 2006).

No ano de 1950, o Brasil sediou a sua primeira Copa do Mundo, tendo como campeã a seleção do Uruguai. A partida, que deu o título à seleção do Uruguai, ficou conhecida como “Maracanazo” devido ao triunfo uruguaio ter acontecido em solo brasileiro, no estádio do Maracanã, na cidade do Rio de Janeiro, diante de 200 mil torcedores que torciam a favor da seleção

brasileira e que ficaram inconsoláveis com a derrota. A capa do Jornal o Globo, de 17 de julho de 1950 (Figura 1), avaliou o fracasso com uma “decepção amarga” em uma “tarde infeliz dos jogadores patricios”.

Figura 1-Capa do Jornal O Globo de 17 de julho de 1950



Fonte: Acervo jornal O GLOBO, 1950.

Com a evolução do futebol no país, o tema ganhou mais destaque nas páginas de jornais. Um exemplo da importância do uso do jornal para o domínio do futebol é a reconstrução da história do jogador Heleno de Freitas, ídolo do clube Botafogo, do Rio de Janeiro, na década de 1940. A história foi mostrada em uma cinebiografia intitulada Heleno, de 2012. Os produtores do filme contaram com poucas fontes de coleta para a pesquisa e utilizaram principalmente relatos jornalísticos da época e fotografias de jornais (BARROS, 2012). Um dos fatos dramatizados no longa-metragem, de acordo com informações retiradas do Acervo O

Estado de S. Paulo e O Globo (2015), foi a transferência do jogador para o futebol argentino, considerada a maior transação do futebol brasileiro naquela época. Apesar da recepção dos argentinos ter sido grande, como mostrado no título de uma matéria do jornal O Globo, publicada no dia 07 de junho de 1948 (Figura 2), a estadia de Heleno no país foi curta, pois retorna em abril de 1949, em uma negociação com o time Vasco da Gama, do Rio de Janeiro.

**Figura 2** - Capa jornal O Globo de 07 de junho de 1948



**Fonte:** Adaptado acervo O Globo

No título e na chamada da notícia ainda é perceptível a grafia da palavra “crack”, que atualmente se escreve “craque”. As mudanças de ortografia estão presentes nos acervos digitais de jornais, são importantes e interferem diretamente na recuperação da informação, portanto a representação da informação nesses sistemas não deve desconsiderar essa característica fundamental.

É notória a expressividade que representa o esporte ao cenário brasileiro. De acordo com Guterman (2009, p.9) o futebol é uma “construção histórica, gerado como parte indissociável dos desdobramentos da vida política e econômica do Brasil. O futebol, se lido corretamente, consegue explicar o Brasil”

Hoje, diferentemente de como ocorreu na sua origem no país, o futebol contempla todos os públicos e engloba diferentes faixas etárias, gêneros e condição social, se tornando uma hegemonia. É sabido que o esporte é movido pela paixão, mas também pode ser compreendido como um negócio, nesse sentido, suas informações são amplamente pesquisadas.

É evidente a riqueza de informações sobre o futebol passíveis de localização nos jornais. No surgimento dos primeiros clubes de futebol no Brasil, ao final do século XIX e até meados da década de 20 do século XX, o jornal ainda era a principal mídia de comunicação e ainda permaneceu sendo predominante como fonte de informação textual até a intensificação do uso da Internet. Com a chegada da Internet, as fontes de informação são ampliadas, reduzindo o público do jornalismo impresso. No entanto, conforme Caldas (2004, p.14), a sua relevância consiste de não ter a pretensão de inventar nem revolucionar nada, e que “leitores de jornal e usuários da Internet tem interesses e curiosidades diferentes”, os leitores dos jornais buscam textos com profundidade, originalidade, interpretativos e analíticos, com suas implicações e possíveis repercussões, posicionados dentro de um contexto mais amplo, baseados em pesquisa e opinião, distintos da instantaneidade e objetividade do jornalismo on-line.

Considerando esse cenário e também sustentando que os acervos digitais de jornais são fontes ininterruptas, que acompanharam o futebol desde o seu início no Brasil, o caderno esportivo, parte integrante da mídia jornalística, se configura como importante diário dos acontecimentos do esporte. Por meio das informações retiradas do jornal é possível recontar a história dos clubes, ou recriar biografias dos ídolos do esporte. Valentim (2006) evidencia a carência de informações voltadas diretamente ao esporte, que contribuam para auxiliar as organizações e os indivíduos do âmbito esportivo a desempenhar suas atividades.

Jorge e Valentim (2015) apontam a importância da organização esportiva. Segundo os autores, a sistemática deve considerar todos os tipos de informação que estão inseridos nesse cenário, independente do suporte e fluxo (formal ou informal). Com tais características, as informações publicadas nos jornais, contam com diferentes olhares, editadas em épocas, estilos e por jornalistas diferentes, com a contribuição de torcedores, colaboradores, jogadores, especialistas, entre outros.

Com advento das TIC e com a oferta dos jornais na internet por meio dos acervos digitais, é justificada a necessidade de direcionar as tecnologias da web para representação das informações, possibilitando assim que os documentos contidos nessas bases sejam encontrados.

Belluzo (2014, p.29) expõe a necessidade atual de informação no esporte, bastante relacionada com as TIC. Para isso, destaca três iniciativas que devem ser investigadas:

- a) Existe a necessidade de acompanhar os avanços tecnológicos e absorver as suas potencialidades;
- b) É importante que se agregue valor a estes conhecimentos e que sejam desenvolvidas novas metodologias para acessar e tornar acessível e compreensível a massa de informação disponibilizada em redes;
- c) Capacidade de migrar serviços convencionais ao novo meio e gerar novos serviços e produtos de informação, bibliotecas digitais e outros meios e suportes que possibilitem a disseminação da informação e a geração do conhecimento.

O jornal, importante fonte de informação para o futebol, acompanha os avanços tecnológicos salientados pela autora Belluzo (2014), tendo migrado de um suporte em papel, para sistemas informatizados, alimentando assim os acervos digitais. No entanto, conforme a autora, além de tornar disponível essa massa documental, é preciso criar metodologias que assegurem o seu acesso.

Para Moreiro González (2011), os sistemas de organização e representação utilizados na atualidade, sugeriram da necessidade de estruturar a web, e essas linguagens foram adaptadas de linguagens documentais pré-existentes.

Mas, para alcançar um nível de estrutura superior, que considere as transformações que a web propiciou e a precisão que esses vocabulários carecem, é preciso aliar às linguagens pré-existentes, como tesouros, com a introdução de ontologias. Para Moreiro González (2011, p.34):

A nova web, na sua caracterização semântica, requer linguagens de representação do conhecimento baseadas em ontologias para localizar e acessar recursos. Linguagens multicamadas que interajam entre modelos e reutilizem recursos com novos modelos de processamento para estabelecer relações de confiança e provas de autenticidade, para garantir aos usuários a disponibilidade de recursos e tornar seu funcionamento transparente.

Para garantir esse funcionamento transparente em que Moreiro González se refere, os sistemas que são baseados em palavras precisam fazer a recuperação de conceitos e os seus contextos de uso. As ontologias podem criar essa ponte entre o conteúdo e o significado e fornecer vocabulário para descrever objetos e a forma como eles se relacionam. (SCHIESSL, 2015). Ao direcionar as ontologias para um domínio, mais específica ela se torna, fazendo com que as ambiguidades daquele contexto sejam tratadas conforme as suas reais finalidades. Nos acervos digitais de jornais, ainda é necessário atentar ao cenário histórico que esses ambientes contemplam. Assim, as mudanças e variações linguísticas ocorridas durante a passagem das décadas precisam ser avaliadas, como no exemplo da palavra “Tropheo/Tropheu”, que teve alteração na sua grafia e hoje se escreve Troféu, ou das palavras “escanteio” e “futebol”, que antigamente adotavam o termo em inglês “*corner*” e “*foot-ball*”.

As ontologias apresentam-se como um importante aliado na organização da informação e sua consequente recuperação. São utilizadas como a forma de atribuir significados mais exatos para

determinadas especialidades, diminuindo assim, sistematicamente, o nível de ambiguidades. Para Moreira González (2011, p.76) as ontologias permitem a “obtenção de respostas precisas a consultas concretas”, passando de um nível de recuperação da informação, para recuperação do conhecimento.

#### **4 CONSIDERAÇÕES FINAIS**

Como discutido, os jornais, devido a sua longa data de existência e cobertura de fatos históricos são importantes fontes de pesquisa. Diante da difusão das Tecnologias da Informação e Comunicação e da necessidade de ampliar o acesso a esses registros, alguns dos principais jornais do país disponibilizam a memória brasileira descrita nas publicações jornalísticas por meio da Internet, possibilitando assim que maior número de usuários possa acessar esses documentos. Esses acervos reúnem muita informação e são importantes bases históricas, considerados patrimônios digitais e, além da história, também são fontes atualizadas, pois continuam sendo alimentadas pelas edições atuais.

Do ponto de vista do futebol, sabido que o jornal impresso no Brasil teve seu surgimento em 1808 e os clubes mais antigos do esporte nasceram em 1900, é possível constatar que a história do futebol brasileiro está contemplada nas páginas dos folhetins. Além disso, o esporte também dispõe de significativo espaço nas edições diárias, os chamados cadernos esportivos. Os textos são trabalhados com profundidade e originalidade, possibilitando que informações detalhadas sobre jogos, jogadores, clubes e eventos sejam encontradas e reconstruídas.

Assim, diante da relevância que os jornais representam como bases históricas, e da disponibilização desses conteúdos em ambientes digitais, evidencia-se a importância do desenvolvimento e atualização de métodos que possibilitem manter as informações dos patrimônios digitais, como os acervos

de jornais, acessíveis aos seus usuários. E quando nos referimos em acesso, precisamos contextualizar os elementos além da digitalização e armazenamento, mas os mecanismos de tratamento para que os documentos sejam encontrados pelos usuários. E com a visão dos acervos digitais, também é preciso considerar a pluralidade dos documentos digitalizados e os nascidos digitais e a transitoriedade do vocabulário inerente aos textos jornalísticos, que testemunham às transformações ocorridas na sociedade e as diferenças socioculturais e geográficas na evolução dos anos.

Assim, sabido que as ontologias não se limitam a estruturas pré-definidas de relações semânticas, mas permitem que essas relações sejam refinadas e incorporadas a sua estrutura de acordo com a necessidade de representação determinada, se aplicadas no contexto dos acervos digitais de jornais, também podem potencializar o acesso às informações.

## REFERÊNCIAS

ALVARENGA, Lídia; SILVA, Daniela Lucas. Organização e representação do conhecimento na ciência da informação: revisão da literatura. **Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 3, n. 1, p. 47-84, jan./dez, 2010. Disponível em: <<http://basessibi.c3sl.ufpr.br/brapci/index.php/article/view/0000009338/11857f0b39c3464c4981dd83bfb51948>>. Acesso em: 09 nov. 2016.

ARELLANO, Miguel Angel. Preservação de documentos digitais. **Ci. Inf.**, Brasília, v. 33, n. 2, p.15-27, 2004. maio/ago. Disponível em: <<http://www.scielo.br/pdf/ci/v33n2/a02v33n2.pdf>>. Acesso em: 28 maio 2016.

ASSIS, Machado de. **O jornal e o livro**. Correio Mercanti. Rio de Janeiro, jan. 1859. Disponível em: <<http://machado.mec.gov.br/images/stories/pdf/cronica/macrl3.pdf>>. Acesso em: 13 ago. 2016.

ANJ. **Linha do tempo com a cronologia da história do jornal no Mundo**. 2008. Disponível em: <<http://www.anj.org.br/cronologia>>. Acesso em: 22 maio 2014.

BARROS, Ernesto. **Craque Heleno de Freitas renasce no cinema**. Jornal do Commercio. São Paulo, p. 1-1. 14 mar. 2012. Disponível em: <<http://jconline.ne10.uol.com.br/canal/cultura/cinema/noticia/2012/03/14/craque-heleno-de-freitas-renasce-no-cinema-35671.php>>. Acesso em: 17 ago. 2016.

BELLUZZO, Regina Celia Baptista. **O acesso e uso da informação no esporte: uma competência cidadã**. 2014. Disponível em: <<http://www.revistacedocseme.prefeitura.sp.gov.br/wp-content/uploads/2014/05/oacessoeousodainformacaooesporteapresentacao.pdf>>. Acesso em: 13 ago. 2016.

CALDAS, Alvaro. **Deu no jornal: o jornalismo impresso na era da Internet**. 2. ed. São Paulo: PUC-Rio; São Paulo: Loyola, 2004. 207 p.

CARRAVETTA, E. S. P. **Modernização da gestão no futebol brasileiro: perspectivas para a qualificação do rendimento competitivo**. Porto Alegre: Age, 2006.

CASTELLS, Manuel. **A galáxia da internet: reflexões sobre a internet, negócios e sociedade**. Rio de Janeiro: Zahar, 2003. 243p.

CAVALCANTE, L. E. . Patrimônio digital e informação: política, cultura e diversidade. **Encontros Bibli**, Florianópolis, v. 23, p. 152-170, 2007. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/viewFile/1518-2924.2007v12n23p152/401>>. Acesso em: 01 ago. 2016.

CHAPOUTHIER, Georges. **Registros evolutivos**. Viver Mente & Cérebro: Memória, São Paulo, n. 2, p. 8-13, jul. 2006.

DINES, Alberto. **O papel do jornal**: uma releitura. 7. ed. São Paulo: Summus, 2001. 157p.

DODEBEI, Vera. **Patrimônio digital virtual**: Herança, documento e informação. Trabalho apresentado na 26ª. Reunião Brasileira de Antropologia, realizada entre os dias 01 e 04 de junho, Porto Seguro, Bahia, Brasil. 2006. Disponível em: <[http://www.abant.org.br/conteudo/ANAIS/CD\\_Virtual\\_26\\_RBA/grupos\\_de\\_trabalho/trabalhos/GT%2037/vera%20dodeber.pdf](http://www.abant.org.br/conteudo/ANAIS/CD_Virtual_26_RBA/grupos_de_trabalho/trabalhos/GT%2037/vera%20dodeber.pdf)> Acesso em: 09 ago. 2016.

DODEBEI, Vera. Cultura digital: novo sentido e significado de documento para a memória social? **DataGramZero**, Rio de Janeiro, v. 12, n. 2, 2011. Disponível em: <[http://www.brapci.ufpr.br/brapci/\\_repositorio/2011/04/pdf\\_15b7c5a842\\_0016272.pdf](http://www.brapci.ufpr.br/brapci/_repositorio/2011/04/pdf_15b7c5a842_0016272.pdf)>. Acesso em: 09 ago. 2016.

FERRARI, Pollyana. **Jornalismo digital**. São Paulo: Contexto, 2006. 120p.

GUTERMAN, Marcos. **O futebol explica o Brasil**: uma história da maior expressão popular do país. São Paulo: Contexto, 2009. 270 p.

JENKINS, Henry. **Cultura da convergência**. São Paulo: Aleph, 2008. 380p.

JORGE, Carlos Francisco Bitencourt. **Gestão da informação esportiva no contexto da inteligência competitiva em clubes de futebol**: um estudo de caso no Marília Atlético Clube. 2013. 322 f. Dissertação (Mestrado) - Curso de Ciência da Informação, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2013. Disponível em: <<http://www.marilia.unesp.br/Home/Pos->

Graduacao/CienciadaInformacao/Dissertacoes/jorge\_cfb\_me\_mar.pdf f>. Acesso em: 13 ago. 2016.

JORGE, Carlos Francisco Bitencourt; VALENTIM, Marta Lígia Pomim. Informação e esporte: a informação esportiva e sua relação com clubes de futebol. **Inf. Inf**, Londrina, v. 1, n. 20, p.183-208, jan. 2015. Trimestral. Disponível em: <[http://www.uel.br/revistas/uel/index.php/informacao/article/view/19712/pdf\\_50](http://www.uel.br/revistas/uel/index.php/informacao/article/view/19712/pdf_50)>. Acesso em: 13 ago. 2016.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Metodologia científica: ciência e conhecimento científico, métodos científicos, teoria, hipóteses e variáveis, metodologia jurídica**. 6. ed. rev. e ampl. São Paulo: Atlas, 2012. 314 p.

LE GOFF, Jacques. **História e memória**. Campinas: UNICAMP, 1990, 553 p.

LE GOFF, Jacques. **Documento/monumento**. In: História e memória. 5. ed. Campinas, SP: UNICAMP, 2003. p. 525-539.

LÉVY, Pierre. **Cibercultura**. São Paulo: Ed. 34, 2000. 260p.

LOPES, Carlos. **A imagem e o sonho da arquivística**. Rio de Janeiro: Arquivo Público, 1998.

MÁXIMO, João. Memórias do futebol brasileiro. Estudos Avançados, [s.l.], v. 13, n. 37, p.179-188, set. 1999. **FapUNIFESP (SciELO)**. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0103-40141999000300009&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0103-40141999000300009&script=sci_arttext)>. Acesso em: 17 ago. 2016.

MONTENEGRO, Rosilene Dias ; SILVA, Fábio Ronaldo da . **Por uma digitalização da memória jornalística**. BOCC. Biblioteca Online de Ciências da Comunicação, Portugal, p. 1-10, 2005. Disponível em: <<http://bocc.ubi.pt/pag/rosilene-montenegro-fabio-silva-memoria-jornalistica.pdf>>. Acesso em: 13 ago. 2016.

MOREIRO GONZÁLEZ, José Antônio. **Linguagens documentárias e vocabulários semânticos para a web**. Salvador: EDUFBA, 2011.

MOURA, Dione Oliveira. O ‘sincrônico’, o ‘diacrônico’, o acontecimento e a errância de sentidos na análise do discurso jornalístico. **Comunicação & Informação**, Brasil, v. 2, n. 12, p.63-73, jul. 2009. Semestral. Disponível em: <<http://www.revistas.ufg.br/index.php/ci/article/view/12270/8132>>. Acesso em: 15 set. 2016.

PALACIOS. Marcos. Ruptura, Continuidade e Potencialização no Jornalismo Online: o Lugar da Memória. In: MACHADO, Elias; PALACIOS, Marcos (Orgs.), **Modelos do Jornalismo Digital**. Salvador: Editora Calandra, 2003. p.1-17.

RUEDA, Valéria Matias da Silva; FREITAS, Aline de; VALLS, Valéria Martin. Memória Institucional: uma revisão de literatura. **Crb-8 Digital**, São Paulo, v. 8, n. 1, p.78-89, abr. 2011. Disponível em: <<http://revista.crb8.org.br/index.php/crb8digital/article/viewFile/62/64>>. Acesso em: 27 jun. 2016.

SANTOS, Luana Carla de Moura dos; VIERA, Angel Freddy Godoy. Avaliação da recuperação da informação em acervos digitais de jornais. **Em Questão**, Porto Alegre, v. 2, n. 21, p.49-73, maio 2015. Disponível em:

<<http://www.seer.ufrgs.br/EmQuestao/article/viewFile/52362/35082>>. Acesso em: 20 mar. 2016.

SCHIESSL, Marcelo. **Lexicalização de Ontologias: o** relacionamento entre conteúdo e significado no contexto da Recuperação da Informação. 2015. 259 f. Tese (Doutorado) - Curso de Ciência da Informação, Programa de Pós-graduação em Ciência da Informação da Universidade, Universidade de Brasília, Brasília, 2015. Disponível em: <[http://www.repositorio.unb.br/bitstream/10482/18663/1/2015\\_MarceloSchiessl.pdf](http://www.repositorio.unb.br/bitstream/10482/18663/1/2015_MarceloSchiessl.pdf)>. Acesso em: 02 mar. 2017.

SCHWINGEL, Carla; ZANOTTI, Carlos A. **Produção e colaboração no jornalismo digital**. Florianópolis: Insular, 2010. 239 p.

TELLAROLI, Taís Marina; ALBINO, João Pedro. **Da sociedade da informação às novas tic's**: questões sobre internet, jornalismo e comunicação de massa. Núcleo de Conforto Ambiental – UNESP. 2007. Disponível em: <<http://www.faac.unesp.br/publicacoes/anaiscomunicacao/textos/28.pdf>>. Acesso em: 14 ago. 2016.

TRIVIÑOS, Augusto N. S. **Introdução à pesquisa em Ciências Sociais**: a pesquisa qualitativa em Educação. São Paulo: Atlas, 2010.

UNESCO. **Carta sobre a preservação do patrimônio digital**. 2003. Disponível em: <<http://unesdoc.unesco.org/images/0013/001311/131178e.pdf>> Acesso em : 10 ago. 2016.

VALENTIM, M. L. P. Processo de inteligência competitiva organizacional. In: VALENTIM, M. L.P. (Org.). **Informação, conhecimento e inteligência organizacional**. 2.ed. Marília: FUNDEPE Editora, 2006. 282p; p.9-24.

VAZ, Paulo. Mediação e tecnologia. **Revista Famecos**, Porto Alegre, v. 1, n. 16, p.45-59, 2001. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/revistafamecos/article/view/3137/2408>>. Acesso em: 06 set 2016.

**DIGITAL NEWSPAPER ACCOUNTS: CONSIDERATIONS REGARDING  
THE RECOVERY OF FOOTBALL INFORMATION**

**Abstract:** *It presents the newspaper as important source of information for historical research of a society. It addresses digital collections, environments that function as on-line newspaper archives as a possibility to access these documents more easily. It discusses how information retrieval should be improved to enable this information to be accessed in a user-friendly way and less dependent on the user's literacy. It points out that this recovery must consider linguistic and sociocultural variations, as well as the specificity of each knowledge domain. To this end, it exposes the importance of the newspaper for football, the shortcomings in the organization of this information in this domain of knowledge and presents the ontology as an efficient knowledge organization system to favor the digital collections of newspapers. By directing the ontologies to a domain, in this case football, the more specific it becomes, making the ambiguities of that context treated according to its real purposes.*

**Keywords:** *Journal digital collection. Soccer. Organization and Information Retrieval.*

*Originais recebidos em: 02/04/2017*

*Aceito para publicação em: 20/06/2017*

*Publicado em: 30/06/2017*