



## **A APURACAO NO JORNALISMO DIGITAL EM BASES DE DADOS.<sup>1</sup>** **(Os casos das plataformas do Kuro5hin, Slashdot e Pacjor)**

Diego Acássio Beal Kerber<sup>2</sup>

**Resumo:** Neste artigo apresentamos os resultados parciais de trabalho feito no Laboratório de Pesquisa em Jornalismo Digital do Mestrado em Jornalismo da UFSC, que busca o estudo conceitual das particularidades do processo de apuração de informações em plataformas jornalísticas, através da revisão da bibliografia especializada sobre jornalismo digital em bases de dados e o estudo das plataformas do Slashdot e Kuro5hin em operação e que incluem a etapa da apuração no sistema. Os subsistemas e aplicativos existentes nestas plataformas são analisados e conceituados, objetivando selecionar os mais adequados à prática do jornalismo digital para incorporação na Plataforma Automatizada para Conteúdos Jornalísticos (PACJOR), em desenvolvimento pela equipe do LAPJOR, com apoio do CNPq.

**Palavras-chave:** plataformas jornalísticas; apuração; bases de dados

### **1. Apresentação**

O uso de bases de dados pelas empresas jornalísticas intensifica-se em meados dos anos 80 (SMITH, 1980) e ao longo deste período estas tecnologias foram sendo aperfeiçoadas para atender as necessidades específicas da prática do jornalismo, passando de funções acessórias de armazenamento de informações a ocupação de funções centrais nos processos de produção e gestão das informações nas redações digitalizadas. Este processo desenvolveu-se com a evolução das bases de dados partindo de sistemas mais genéricos de organização de informações para sistemas mais específicos para o jornalismo, buscando a otimização das ações do jornalista nos processos de produção, recuperação, gerenciamento e circulação das informações. Barbosa (BARBOSA, 2007, p. 152) aponta as bases de dados como o paradigma de uma nova etapa no jornalismo digital. Nesta etapa, ainda em desenvolvimento, através do uso das bases de dados a estruturação da informação ocorre de forma automatizada, aumentando o potencial de inovação e flexibilidade nas relações entre jornalistas, fontes e membros do público, o que vêm possibilitando o aparecimento de formas diferenciadas de utilização do computador como ambiente de trabalho no jornalismo.

As plataformas automatizadas para a produção de conteúdos são estruturas complexas articuladas em torno de tecnologias, softwares e aplicações informáticas

---

<sup>1</sup> Trabalho apresentado ao Intercom Junior, na Divisão Temática de Jornalismo, do X Congresso de Ciências da Comunicação

<sup>2</sup> Graduando em Jornalismo na Universidade Federal de Santa Catarina (UFSC). Membro do Laboratório de Pesquisa Aplicada em Jornalismo (LAPJOR). Bolsista PIBIC/CNPq.



pensadas para, através de processos padronizados em consonância com os critérios utilizados pelo jornalismo, auxiliar o jornalista nas suas funções (SCHWINGEL, 2008). Estas plataformas mais complexas são desenvolvidas para explorar as capacidades de construção de narrativas típicas desta forma de jornalismo, com características como multimídia, hipertextualidade e interatividade. (SALAVERRÍA, 2005, p. 25-34) Ao mesmo tempo em que são utilizadas na construção das narrativas, as plataformas automatizadas armazenam de forma organizada as informações. Estruturadas na forma de bases de dados, as plataformas de produção de conteúdos jornalísticos podem operar na organização do fluxo de produção, auxiliando na coordenação das atividades dos jornalistas que atuam no processo de produção e gestão de informações jornalísticas.

Para que o trabalho jornalístico possa ser realizado de forma completa, uma plataforma deve incluir todas as etapas da produção, desde a apuração, passando pela produção, até a circulação. No caso do ciberjornalismo a digitalização permeia todas as três etapas (MACHADO, 2006, BARBOSA, 2007). Do ciberespaço o jornalista retira as fontes e as informações que incluirá na matéria, na plataforma irá compor a notícia, inserindo texto e os elementos multimídia que o subsistema de publicação possibilita agregar, e a publicação ocorrerá no próprio ciberespaço, onde todos os membros da rede existente em torno do cibermeio poderão ter acesso ao conteúdo produzido, podendo contribuir para sua modificação ou mesmo para intensificar a sua circulação. (SALAVERRÍA, 2005, p.21)

Ao longo de mais de uma década identificamos vários modelos de plataformas para a produção de conteúdos jornalísticos. Desde os mais simples até os mais complexos, a depender da diversidade das funções existentes e da integração das três etapas: apuração, produção e circulação. As plataformas menos complexas estão centradas na produção, como nos casos de projetos experimentais como *OpenNews* desenvolvida pelo *Laboratório Kimera*, da Universidade Federal de Pernambuco e do *Panopticon*, criada pela equipe do GJOL da FACOM/UFBA. (MACHADO; PALACIOS, 2007). Neste artigo, em primeiro lugar, focarei minha atenção no estudo das particularidades dos processos de apuração nas plataformas que auxiliam o jornalista nesta etapa de produção dos conteúdos jornalísticos. Com este propósito selecionamos as plataformas do *Kuro5hin* (<http://www.kuro5hin.org/>) e do *Slashdot* (<http://slashdot.org/>). Em segundo lugar, apresento alguns detalhes do subsistema de apuração da Plataforma Automatizada para Conteúdos Jornalísticos (PACJOR), em desenvolvimento no LAPJOR/UFSC.



## 2. Atuação das plataformas na apuração jornalística

Uma plataforma projetada para atuar na etapa de apuração deve auxiliar o repórter na obtenção de informação. Como no caso destas plataformas tratamos do jornalista digital, ou do ciberjornalista, o espaço de investigação e recolhimento das informações é o próprio ciberespaço (MACHADO, 2003). Com este novo campo de ação, muda a forma como o jornalista apura as informações. (QUINN, 2001) Nesta situação o profissional trabalha muito mais verificando as informações que foram disponibilizadas na internet com maior relevância para agregar a uma matéria. Este processo diferencia-se da antiga forma de apuração. Agora o processo de produção de informações está muito mais descentralizado. O jornalista, ainda que mantenha as antigas funções de apuração, atua mais como um editor, verificando as informações disponibilizadas na *web* que atendem aos critérios de noticiabilidade e são reputadas como importantes pelos membros das redes de circulação de informações. Neste processo, ele deverá buscar o máximo de diversidade nas informações, além de possibilitar uma maior contextualização da notícia publicada. (BRUNS, 2005, p.17)

A mudança na logística do processo de produção de informações jornalísticas abre a possibilidade alterações relevantes na forma como são produzidas as notícias. Agora qualquer pessoa com acesso à internet e com um sítio web pessoal, mesmo sem formação jornalística, pode publicar informações que julga importantes sem passar pela filtragem de jornalistas, procedimento antes sem espaço no modelo convencional de jornalismo centralizado (ALAN, 2006). Simultaneamente, aumenta o número de informações sem relevância do ponto de vista jornalístico, o que demanda a criação de mecanismos profissionais e automatizados, capazes, por um lado, de operar a filtragem dos conteúdos publicados e, por outro, aproveitar a produção descentralizada.

As plataformas analisadas neste artigo, *Slashdot* e *Kuro5hin* possibilitam a inserção de conteúdos por todos as pessoas cadastrados no sistema, agregando grande quantidade de informações. Este conteúdo enviado é o ponto de partida para a criação de uma matéria. Assim a apuração é realizada por todas as pessoas cadastradas, que através do *gatewatching* podem localizar mais informações relacionadas e complementar as notícias publicadas. Nenhuma destas plataformas trabalha com o conceito de pauta como no caso da proposta implementada a partir da Plataforma Panopticon no Educação em Pauta e prevista para ser incorporada numa versão



atualizada da Panopticon que acabou não sendo desenvolvida. (SCHWINGEL, 2007; MACHADO; PALACIOS, ROCHA; SCHWINGEL, 2007) A pauta jornalística atua como forma de contextualização dos fatos para o repórter e indica o direcionamento dos aspectos que devem ser abordados em cada matéria. Na pauta estão contidas informações iniciais que situam o repórter acerca do tema abordado, além de indicar um enfoque a ser que deve possuir o texto e apresentar sugestões de fontes que devem ser entrevistadas para obtenção de mais informações. (LAGE, 2001)

Nos casos aqui analisados, os usuários cadastrados postam diretamente as propostas de notícias na plataforma, sem a presença de uma pauta, e a publicação ou mesmo a alteração dos conteúdos vai depender do sistema de edição utilizado, podendo ser mais ou menos centralizado. No *Slashdot* todos podem postar as matérias, mas o material recebido é editado apenas por um corpo editorial selecionado pela equipe plataforma, de acordo com critérios previamente estabelecidos. A adição de novas informações ou a refutação do material apresentado só pode ser feito pelos demais usuários através da seção de comentários, sem que exista a possibilidade de agregar estas informações ao corpo principal da matéria. (BRUNS, 2005) Já no *Kuro5hin* mais pessoas podem editar as informações e participar da apuração, adicionando mais conteúdo à narrativa ou alertando sobre a falta de veracidade em trechos das matérias. Este modelo é mais interessante por possibilitar maior versatilidade nas interações entre as pessoas que participam no desenvolvimento de um conteúdo, além de permitir que informações relevantes sejam adicionadas ao corpo principal da notícia, não ficando relegadas aos comentários (BRUNS, 2005; MACHADO, 2008).

Em nenhuma destas plataformas existe um cadastro geral de fontes (CORREIA;SCHWINGEL, 2008). As duas plataformas são estruturadas considerando que a maior parte do conteúdo será obtida através do *gatematching*, ou seja, estes conteúdos já estarão disponíveis e, portanto, dispensam-se as entrevistas realizadas diretamente com as fontes, como ocorre no jornalismo em outros suportes. Como as plataformas de produção de conteúdos jornalísticos armazenam uma quantidade considerável de informação, a partir do momento que possuem elevado número de pessoas inserindo conteúdos ou interagindo com as informações produzidas pela comunidade cadastrada, é necessário criar um filtro automatizado para identificar, por exemplo, o que é mais relevante para se utilizar numa matéria, ou o que pode ser o ponto de partida para uma nova apuração. Neste caso a plataforma deverá usar critérios mais próximos possíveis dos que seriam aplicados por um jornalista. Dentre os



mecanismos presentes nas plataformas estudadas, o *Slashdot* apresenta um dos mais inovadores, que é o uso do *Karma* (CHAN, 2002; BRUNS, 2005, p.45).

O foco do *Karma* é servir como a avaliação das contribuições dos colaboradores cadastrados, de acordo com avaliações atribuídas pelos demais colaboradores. Esta avaliação consiste em categorizar um comentário como ruim, bom ou moderado. O colaborador que possui um “karma” positivo receberia funções de maior importância no processo de produção das informações, assim como os conteúdos de sua autoria ganham mais destaque. Um próximo passo, para aproveitar ainda mais esta informação armazenada, seria transformar esta pessoa com contribuições diferenciadas em uma fonte, nos casos em que a matéria tem como assunto algo da especialidade deste usuário. Outros conteúdos, como é o caso dos comentários, também passam por um processo de seleção, com o intuito de destacar os mais relevantes. No caso do *Kuro5hin*, o próprio colaborador pode selecionar um filtro para os comentários que devem ser exibidos. Assim o leitor pode ver todos os comentários na ordem que foram postados, ou apenas os últimos e com melhores cotações pelos outros internautas. (BRUNS, 2007, p. 45-47)

Cada uma das plataformas apresenta modelos diferentes para a definição dos conteúdos que são exibidos na página inicial e que, portanto, tem maior destaque no sítio web. O *Slashdot* adota um modelo mais fechado de produção, pois somente um corpo delimitado de editores faz esta seleção dos conteúdos a serem exibidos. (BRUNS, 2007, p. 38-40) Para uma plataforma que busca a máxima integração da maior parte das pessoas articuladas em torno de cada cibermeio, este modelo se configura como tendo uma certa limitação, pois exclui da edição a maioria das pessoas que participam da produção dos conteúdos.

De forma diferente opera a plataforma do *Kur5hin*, que atua utilizando o conceito de edição aberta. Neste sistema todos os usuários participam da edição de uma matéria, desde o momento em que foi postada a história, até a seleção do grau de importância que receberá na publicação, sendo como capa do sítio web ou apenas de uma editoria específica. (BRUNS, 2005, p.157) Este modelo é mais eficiente, pois torna viável o auxílio de todos os membros cadastrados na produção de um conteúdo, seja com opiniões pessoais acerca do encaminhamento da produção ou com a inserção de novas informações.

Outro recurso interessante que possibilitaria um melhor aproveitamento das diversas informações armazenadas no banco de dados destas plataformas é o *data*



*minning*. A mineração de dados, como define o autor, é a habilidade de tratar grandes quantidades de dados, e através deste processo adquirir novas informações, que de outra forma permaneceriam inacessíveis. (COLLE, 2002, p.23) Desta forma, pode-se encontrar informações úteis que podem servir de ponto de partida para uma nova matéria. As possibilidades são múltiplas, e podem incluir a criação de uma pauta a partir das palavras-chave mais buscadas através do mecanismo de busca, ou da notícia mais comentada recentemente, ou das *tags* presentes nas notícias mais acessadas, etc. Para tornar este *data minning* possível, estas plataformas devem incluir meta-dados, neste caso representado pelas tags. Os meta-dados são informações acerca das informações. Servem como etiquetas que classificam o tipo de informação que consta nesta matéria. (COLLE, 2002, p. 33)

### **3. Uma proposta para atuação de uma plataforma no processo de apuração**

A Plataforma Automatizada para Conteúdos Jornalísticos (PACJOR) em desenvolvimento no Laboratório de Pesquisa em Jornalismo Digital (LAPJOR), tem como objetivo auxiliar na produção de conteúdos jornalísticos sobre o Centro de Comunicação e Expressão da Universidade Federal de Santa Catarina (UFSC) e está sendo projetada para ser utilizada por estudantes da graduação do Curso de Jornalismo da Universidade e pelos membros da comunidade do CCE. As principais diferenças existentes se comparada às plataformas até aqui discutidas neste artigo, estão na articulação entre todas as etapas da produção de informações (apuração, produção e circulação) e a inclusão no sistema de apuração de um subsistema para fontes e para pautas, pois prevê que os participantes cadastrados no sistema usarão além do *gatematching*, os demais métodos de apuração adotados pelo jornalista profissional, como a entrevista, a pesquisa documental, etc. A PACJOR será estruturada de forma a que possua um subsistema de fontes e um subsistema de pautas. (SCHWINGEL; CORREIA, 2008)

#### **3.1 O subsistema de fontes**

O subsistema de fontes operará através do armazenamento das informações acerca das fontes em uma tabela, e seu acesso será disponível a todos os repórteres a qualquer momento da produção. Todos que operam a plataforma terão liberdade para incluir novas fontes, sendo elas pessoas, sítios web interessantes ou documentos. Este



subsistema também realizará a inclusão de usuários cadastrados como fontes em potencial. A seleção de quais serão utilizados em uma matéria se dará através de dados coletados durante o cadastro através de um formulário rápido de seleção de áreas de interesse desta pessoa, e caso o sistema encontre compatibilidades entre o assunto abordado na matéria e este usuário, ele será incluso como fonte em potencial.

O repórter, após a entrevista com a fonte, poderá cotá-la de acordo com o grau de contribuição à matéria. O próprio sistema automaticamente pode, através de um ranking indicar fontes com melhores participações ao longo da produção dos conteúdos da plataforma naquela área ou tema, ordenando as fontes pela pontuação em proporção ao número e qualidade de contribuições. Desta forma a plataforma será capaz de incorporar o maior número de fontes disponíveis no ciberespaço e, de acordo com os conhecimentos que cada um dos internautas possui os incorporar como fonte em futuras pautas.

#### **Subsistema de registro de fontes**

- Cadastro de fontes oficiais, pessoas, sítio web ou documental, realizado pelos repórteres
- Fontes cadastradas para comentar, através de dados fornecidos no cadastro

### **3.2 O subsistema de pautas**

Visando o uso desta plataforma por estudantes de jornalismo, é importante que o sistema trabalhe com o conceito de pauta. A pauta auxilia na produção da notícia, situando o repórter acerca do assunto que irá investigar, além de possibilitar um maior controle das atividades que cada repórter está realizando. (LAGE, 2001)

Nesta pauta devem estar discriminados diversos tópicos, como o tema a ser abordado na matéria, um histórico com informações relacionadas, um enfoque que servirá de norte para a narrativa e trará um hipótese a ser estudada, e uma seleção de fontes a serem consultadas, sendo elas tanto pessoas quanto documentais. (MACHADO; PALACIOS, ROCHA; SCHWINGEL, 2007, p.123)

Além dos repórteres que atuam na plataforma, pessoas que acessam a plataforma para visualizar as notícias ou comentar, devem ser motivadas a sugerirem pautas. Estas sugestões de pauta tornam-se pautas de acordo com uma cotação, atribuída por



jornalistas que utilizam a plataforma e pelos internautas. (BRUNS, 2005, p.157)

A cotação é atribuída pela seleção das opções positivo ou negativo, e considera-se uma pauta potencial aquela com um número expressivo de cotações, este número ponderado de acordo com o número de acessos, sendo estas cotações em sua maioria positiva. As com maioria negativa são descartadas devido sua baixa aceitação, e as com poucas cotações também devem ser desconsideradas, pois a sugestão de pauta não despertou interesse, logo não atendeu a um critério importante de noticiabilidade.

É importante possibilitar a participação das pessoas que acessam o sítio web para integrá-los ao processo, e dotá-los de uma parcela de controle do sistema de produção de informações, de tal forma que desenvolvam maior interesse em participar na produção, graças à noção de pertencimento que desenvolvem neste processo. (BARBOSA, 2007)

Assim que a pauta alcançou um dos critérios requeridos pela plataforma para ser aceita, ela entra na fila de espera para apuração. Cada repórter poderá selecionar uma pauta nesta fila, e iniciar os trabalhos para a produção da matéria. Neste processo os demais usuários poderão ver a pauta e qual repórter está encarregado da produção da matéria, e poderão fazer comentários sobre a pauta, possibilitando uma troca de informações entre todos os usuários.

O sistema auxiliará na criação da pauta, através das *tags* inclusas. Através delas, ele apontará matérias com *tags* coincidentes, ordenadas por data, de todos os conteúdos já produzidos. Com o uso adequado das *tags*, a possibilidade de apontar conteúdos anteriores armazenados relevantes para uma nova pauta é grande, e a plataforma desta forma atuando também como um banco de memória, e trazendo a tona informações do passado e realizando a relação com os fatos presentes e repercussões futuras. (MACHADO, 2006, p. 27)

Desta forma, o sistema também pode indicar pautas sendo apuradas por outro repórter com tema semelhante, melhorando a organização das atividades, tornando mais viável uma troca de informações entre os repórteres com assuntos relacionados e evitando que ambos trabalhem para obter as mesmas informações. Desta forma a plataforma atua gerindo toda a produção da equipe envolvida na produção e maximiza o aproveitamento dos recursos humanos, tornando o trabalho mais coletivo. (MACHADO, 2006, p. 33) As *tags* também viabilizam a indicação de fontes, através do subsistema de cadastro de fontes, realizando uma comparação entre as utilizadas em outras matérias com *tags* coincidentes, logo com uma área de conhecimento semelhante.



Assim estruturada, a plataforma do PACJOR auxiliará no controle da produção de todos os repórteres, situa-os na cadeia de produção e disponibiliza todas as informações armazenadas na base de dados sobre o assunto da pauta que trabalham, sendo estas informações já cristalizadas em forma de uma notícia divulgada, ou na forma de uma sugestão de pauta, pauta ou matéria não editada, logo a maior parte das informações disponíveis na intranet. Construída desta forma, a plataforma facilita o compartilhamento de informações, e motiva uma atitude chave de uma *intranet*: compartilhar a informação e assim agregar mais valor aos conteúdos. (QUINN, 2001, p. 98)

O sistema também poderá fazer suas sugestões de pauta através do *data mining* dos dados armazenados nas tabelas. A plataforma pode sugerir para ser apurada uma matéria com a palavra mais buscada no mecanismo de busca como tema, ou indicar como tema a ser investigado uma notícia com grande número de comentários, com o objetivo de escrever uma matéria da repercussão do assunto. Isto possibilita um melhor aproveitamento de informações armazenadas na própria plataforma.

#### **Subsistema de pautas**

- Cadastro de sugestões de pauta, realizado por todos os usuários
- Exibição de pautas ou matérias relacionadas por *tags*, armazenadas no sistema
- Sugestão de fontes através de comparação de *tags*
- Sugestão de pautas pelo próprio sistema, através do *data mining*
- Fila de aprovação de pautas

#### **4. Conclusão**

A contribuição das plataformas do *Kuro5hin* e *Slashdot* para o jornalismo são inegáveis. Elas representam o processo de abertura pelo qual o jornalismo passa neste momento, e reconhecem que alguns procedimentos técnicos atribuídos aos jornalistas como entrevistar, apurar ou editar já não são atividades exclusivas dos profissionais, e que os conteúdos adicionados por estes indivíduos podem complementar a produção de jornalistas, enriquecendo-as. (CUNHA, 2007, p.102)



Estas plataformas são adaptadas ao uso do *gatewatching*, e toda a logística de apuração que este método possui, porém não apresentam subsistemas de apuração pensados na forma como o jornalismo profissional é praticado, ignorando conceitos importantes, como por exemplo, a pauta.

Desta forma, para o desenvolvimento de uma plataforma que realize com o máximo de eficiência os objetivos propostos pelo LAPJOR, subsistemas de apuração deverão complementar o funcionamento desta plataforma, em comparação com as apresentadas neste artigo. Este subsistema deve valorizar ao máximo o recolhimento de informações e organização dos repórteres, aumentando desta forma a qualidade e eficiência da produção jornalística.

A forma mais adequada para alcançar este objetivo é ao mesmo tempo em que se incorporam os métodos do *gatewatching* criados pelas comunidades de redes sociais para utilizar o ciberespaço como campo de investigação, utilizar os métodos já consagrados pela prática ao longo dos anos pelo exercício do jornalismo profissional.

Plataformas com a inclusão de subsistemas de apuração já foram estudadas e desenvolvidas. Neste caso enquadramos sistemas como o *Panopticon*, desenvolvido pela equipe de pesquisadores do Grupo de Jornalismo On-line, do Programa de Pós-Graduação em Comunicação da Universidade Federal da Bahia. Nele já observamos a presença de um subsistema atuante na produção de pautas. (MACHADO; PALACIOS, ROCHA; SCHWINGEL, 2007, p. 122-123) O diferencial que buscamos ao longo do desenvolvimento do PACJOR é uma integração mais completa entre todas as etapas de produção de conteúdos jornalísticos, unindo da forma mais eficiente possível a apuração, produção e circulação das informações produzidas, em um sistema que agregue mais conhecimento conceitual e prático, e avance no sentido de criar plataformas desenvolvidas em bases de dados que atendam de forma mais adequada às especificidades do jornalismo digital.

## 5. Bibliografia

BARBOSA, Suzana. **Jornalismo Digital em Bases de Dados. Um paradigma para produtos jornalísticos digitais dinâmicos**. Salvador: Universidade Federal da Bahia, 2007

BRUNS, Axel. **Gatewatching. Collaborative online news production**. London: Peter Lang, 2005.



COLLE, Raymond. **Explotar información noticiosa: Data minig aplicado a la documentación noticiosa.** Madrid: Universidad Complutense de Madrid, 2002.

CUNHA FILHO, Paulo. **Da definição da forma-jornal à construção do OpenNews: Um sistema para o jornalismo de intervenção** in MACHADO, Elias e PALACIOS, Marcos. **O Ensino de Jornalismo em Redes de Alta Velocidade. Metodologias e Softwares.** Salvador: EDUFBA: 2007.

MACHADO, Elias e PALACIOS, Marcos. **O Ensino de Jornalismo em Redes de Alta Velocidade. Metodologias e Softwares.** Salvador: EDUFBA: 2007.

MACHADO, Elias. **O Jornalismo Digital em Base de Dados.** Florianópolis: Calandra, 2006.

MACHADO, Elias. **O ciberespaço como fonte para os jornalistas.** Salvador. Calandra. 2003.

MANOVICH, Lev. **The language of new media.** Cambridge: MIT Press, 2001.

QUINN, Stephen. **Convergent Journalism: The Fundamentals of Multimedia Reporting.** London. Peter Publishing, 2005. media. Baltimore: The Johns Hopkins University Press, 2001.

SALAVERRIA, Ramón. **Redacción periodística en internet.** Pamplona: Ediciones Universidad de Navarra, S. A., 2005.