

PERSPECTIVAS

Artigo convidado

Versão original

DOI: <http://dx.doi.org/10.1590/S0034-759020180208>

O GATO QUE COPIA: SIMILARIDADES E PRODUÇÃO ACADÊMICA NA ERA DIGITAL

INTRODUÇÃO

Minha experiência como editor da *RAE* permitiu aprendizado sobre as dificuldades de lidar com o tema da similaridade da publicação da produção científica na era digital. Constantemente tinha que arbitrar sobre questões delicadas, e diversas vezes tive a oportunidade de me expressar, em apresentações orais e em editoriais da própria revista, sobre como enfrentar esse tema tão complexo. Pretendo reproduzir, neste curto artigo da seção Perspectivas conceitos que fui adquirindo ao longo daquele período, utilizando material de alguns dos editoriais que escrevi para ajudar a compreender o tema.

Vou evitar, sempre que possível, a expressão “plágio” nesta discussão, que entendo ter o objetivo de informar e educar os autores. Plágio pressupõe dolo, e, na maioria dos casos que enfrentei, o principal problema era de erros de interpretação e falta de entendimento do processo editorial acadêmico. Nos casos onde há má intenção explícita, o problema é de outra ordem, e não vou discuti-los aqui, até porque eles acontecem com menos frequência e devem ser tratados de acordo com preceitos jurídicos, o que não é minha especialidade. O objetivo deste artigo é ajudar os autores bem-intencionados nas diversas situações que eles podem enfrentar ao preparar seus artigos para publicação em periódicos de alto nível.

Em primeiro lugar, para uma boa ilustração da complexidade do problema da similaridade e da diversidade de situações em que ela se manifesta, recomendo consultar o infográfico “Did I plagiarize?” (*The Visual Communication Guy, 2014*). Esse infográfico apresenta uma tipologia caracterizando 13 situações diferentes em que um autor pode incorrer em algum grau de similaridade em seu artigo, situações que identificam a gradação da gravidade do problema enfrentado. Respondendo às perguntas do infográfico, os autores podem aprender muito sobre os limites da reutilização de textos em trabalhos científicos.

Em segundo lugar, estamos em processo de adaptação cultural a um mundo cada vez mais digitalizado. Há até alguns anos, copiar um texto de um lugar para outro era tão ou até mais trabalhoso do que escrever um original. O *copy & paste* é relativamente recente nas nossas vidas, por uma questão técnica, introduzida a partir da disseminação das tecnologias digitais. Como é uma prática fácil e conveniente, disseminou-se rapidamente, e ainda não assimilamos amplamente os riscos embutidos nela.

Em terceiro lugar, e aí mora o problema mais complicado, nós, acadêmicos, sofremos imensa pressão para publicar uma quantidade cada vez maior de artigos. O mais perverso é que

EDUARDO H. DINIZ

eduardo.diniz@fgv.br

ORCID: 0000-0002-7950-9146

Fundação Getúlio Vargas, Escola
de Administração de Empresas
de São Paulo, São Paulo, SP,
Brasil

os sistemas de avaliação da produção científica estão muito bem desenhados para computar a “quantidade” produzida, embora haja muito menos consenso sobre como identificar a “qualidade” dessa produção. É daí que nasce o chamado “produtivismo”, comportamento pernicioso para a evolução do conhecimento científico e que tem sido tema de constante debate na academia, gerando ótimos artigos, inclusive na *RAE* (Machado & Bianchetti, 2011).

Feito este preâmbulo, vamos avançar na discussão sobre as similaridades em duas frentes. A primeira é como tratar diferenças e semelhanças em artigos inteiros, ou seja, como considerar dois artigos que são, na verdade, um único ou um artigo que foi dividido em dois. A segunda é como tratar similaridades encontradas em trechos de artigos, ou seja, material que se retira de um artigo para ser usado em outro.

SOBRE REPUBLICAÇÃO E SALAMES

Inicialmente vou recuperar uma discussão apresentada no editorial “Ética e bom senso contra o produtivismo” (Diniz, 2013), escrito em resposta a uma consulta feita à redação sobre se a revista aceitaria uma segunda versão de artigo anteriormente publicado, seja em outra língua ou com o texto modificado.

Embora a resposta óbvia devesse ser um não, a pergunta me levou a refletir mais profundamente sobre a situação de um autor hipotético que, tendo seu artigo originalmente publicado em chinês, por exemplo, gostaria de ampliar o espaço de divulgação de suas ideias para outros públicos, por exemplo, no Brasil. De alguma forma, não se pode negar que o artigo seria original em português, o que poderia justificar a sua publicação. Por outro lado, ao aceitar esse artigo, a revista poderia estar se desprestigiando ao publicar um artigo não mais inédito, ainda que numa língua nada familiar à grande maioria de seus leitores. Embora essa possa parecer uma questão um tanto bizantina, ela ganha relevância se considerarmos os calorosos debates na nossa academia sobre que língua devem adotar os periódicos nacionais.

Entretanto, o editorial avança em outra direção e coloca uma questão diferente: A quem uma revista pretende atender com a republicação de um artigo? Se a consulta feita pelo hipotético autor diz respeito à sua necessidade de publicação em um periódico mais relevante, pelo fato de a primeira versão em chinês ter acontecido em um periódico de pontuação mais baixa, então claramente estamos diante de uma lógica produtivista. Por outro lado, se o artigo for identificado explicitamente como uma reprodução, evitando que possa ser computado duplamente para

o autor como se fosse uma nova produção, e for avaliado pela editoria que há interesse para a comunidade na qual o periódico está inserido na disseminação de nova versão de um artigo já publicado, então é possível aceitar a ideia de uma republicação.

O importante nesse caso é entender que o compromisso maior do periódico deve ser sempre com seus leitores, e não com a lógica do produtivismo. Embora tenham compromissos claros com os autores que publicam, os periódicos científicos devem sustentar o compromisso fundamental com seus leitores e com os princípios da disseminação do conhecimento de maneira ética. Assim, a melhor resposta à consulta feita à redação deveria ser um grande “depende”. Publicar para atender à necessidade do autor hipotético: Não rotundo! Republicar poderia ser considerado, desde que comunicando aos leitores a existência da publicação original e ressaltando a relevância de ampliar a disseminação de conhecimento relevante, e de maneira ética.

Situação semelhante poderia acontecer com a publicação de uma versão modificada de outro artigo. Seria possível extrair mais de um artigo a partir de investigações empíricas diferentes feitas sobre uma mesma base teórica? Ou, ainda, publicar discussões diferentes feitas sobre uma mesma base empírica? Mais uma vez, temos que isolar a lógica do produtivismo e avaliar as possíveis contribuições dos artigos para a comunidade de leitores a quem o periódico deve servir prioritariamente. Se a potencial contribuição de um artigo se repete no seu similar, não há interesse na publicação da segunda versão. Entretanto, se a introdução de uma nova leitura conceitual sobre uma mesma base de dados ou a aplicação de uma mesma base conceitual sobre diferentes estudos empíricos for capaz de produzir uma nova contribuição ao conhecimento de uma determinada área, a segunda versão pode ser bem-vinda.

Evitando o que é costumeiramente chamado de “pesquisa salame”, aquela que fatia os seus produtos com o único objetivo de produzir um número maior de artigos, é perfeitamente aceitável imaginar que uma mesma pesquisa possa dar origem a diferentes artigos. O importante, mais uma vez, não é a quantidade de artigos, mas a diversidade das contribuições. Ainda que qualquer editor possa preferir uma única contribuição mais completa num único artigo a duas parciais em artigos diferentes, mais uma vez, o que deve ser evitado é a sombra do produtivismo.

Identificar se dois artigos são, na verdade, um único nem sempre é tarefa elementar. Muitos gostariam de exigir que os periódicos explicitassem em suas normas de publicação quais as condições para aceitação ou não de versões modificadas. Entretanto, como foi visto acima, nem tudo é tão preto ou branco. Embora os periódicos tenham a obrigação de explicitar com o máximo de clareza os limites do que é aceitável dentro da sua

linha editorial, cada caso é um caso. E, como conclui o referido editorial, “mais do que qualquer regra, ética e bom senso são o melhor remédio contra o produtivismo”.

LIMITES DAS FERRAMENTAS DE IDENTIFICAÇÃO DE SIMILARIDADE

Num outro editorial (Diniz, 2015b), o tema foi a implantação de uma nova ferramenta de identificação de similaridade na RAE. Instrumento cada vez mais comum nas redações dos periódicos científicos, cheguei a nominar o uso desses sistemas como o “segundo estágio da revolução digital nos periódicos científicos” (Diniz, 2015a).

Motivados pela disseminação da digitalização, periódicos viram-se pressionados a desenvolver estratégias para identificação de similaridades em artigos que recebem. Embora a questão ética sobre integridade do conhecimento que está sendo disseminado pelo periódico seja um ponto relevante na decisão pela adoção dessas ferramentas, não se pode desprezar a sua utilidade para lidar com a questão do direito à propriedade intelectual no mundo acadêmico. Esse segundo ponto – propriedade intelectual – costuma ser, em geral, muito menos óbvio para os pesquisadores do que o primeiro – ética científica.

Ao passar a operar com uma ferramenta de identificação de similaridades, os periódicos adentram questões novas e passam a aprender a diferenciar situações que não eram tão evidentes anteriormente. O que aprendi, com a implantação do sistema na RAE, foi que classificar e tratar os diferentes casos de similaridade se tornou uma atividade bem mais complexa e, portanto, responsável pela ampliação do nível de atrito entre editoria e autores.

A ideia geral que temos sobre esses sistemas é que seu objetivo é evitar que os periódicos publiquem trechos de textos de terceiros sem referenciar o original. Embora negativos, esses casos não são muito problemáticos, pois os autores costumam ser treinados para evitar esse tipo de situação. No caso de terem seus artigos identificados com esse problema, autores, em geral, reconhecem o deslize e costumam aceitar passivamente a solicitação de alteração, refazendo a citação para torná-la mais adequada ao modo reconhecidamente mais ético de se mencionar o trabalho de outros.

Sobre esse tipo de problema de citações incorretas, incompletas ou equivocadas, em outro editorial (Diniz, 2014), chamei a atenção para um artigo da revista inglesa Times Higher Education que menciona o episódio em que o respeitado

intelectual polonês Zigmunt Bauman, criador do conceito de “modernidade líquida”, reagiu a uma acusação de plágio, baseado em denúncia feita por estudante de doutorado (Jump, 2014). De acordo com o editor daquela revista, em sua resposta à acusação, Bauman entende que “intelectuais de prestígio não devem obediência às regras técnicas de referência”, uma opinião não apoiada pela maioria das revistas e, provavelmente, pela comunidade científica no geral. Embora Bauman também tenha deixado claro que ele “nunca deixou de reconhecer a autoria das ideias ou conceitos que inspiraram as que ele mesmo criou”, este episódio não teria acontecido sem a existência de ferramentas de detecção de similaridades.

Se mesmo nesse caso mais simples de identificação de similaridade pode existir alguma dificuldade, há outros casos que têm poder de amplificar o atrito entre periódicos e autores. É muito comum, por exemplo, a similaridade ser identificada em outros textos do mesmo autor. Autores geralmente reutilizam partes de seus trabalhos anteriores, prática conhecida por “autoplágio”, e não costumam ver problema nisso, afinal o texto reutilizado é deles mesmos. Da perspectiva dos periódicos, entretanto, a situação não é tão simples.

Dependendo do texto de origem, a inocente, e despida de más intenções, prática de o autor reutilizar partes de seu próprio material pode gerar problemas complicados para o periódico. Em casos menos graves, quando o texto original não envolve cessão de direitos do autor para a entidade que o publicou, um congresso, por exemplo, muitos periódicos relevam o grau de similaridade e entendem que as versões anteriores publicadas são apenas estágios naturais na elaboração de um artigo.

Tudo fica muito diferente quando o texto de origem envolve um contrato de cessão de direitos, caso da maioria dos periódicos. Nesses casos, o contrato assinado entre autor e entidade publicadora normalmente limita o seu uso, no todo ou em parte, por outra publicação. Assim, quando uma similaridade desse tipo é encontrada, o periódico tem obrigação de solicitar que o autor refaça os trechos similares, sob pena de sofrer um processo por apropriação indevida. Autores que têm dificuldade de entender a separação entre “autoria” e “propriedade” de uma obra intelectual se sentem ofendidos, por acreditarem que não estavam fazendo nada de errado. No entanto, ao assinarem contrato com a editora original, a questão deixa de ser meramente ética e passa a ser do direito comercial.

E a sutileza dos problemas com a identificação de similaridade não para aí. E se o caso for de similaridade em artigos diferentes publicados por uma mesma editora? Tomemos como exemplo este artigo que você está lendo agora. Eu me propus a escrever sobre um tema que já havia tratado anteriormente

em editoriais da *RAE* e anunciei isso no primeiro parágrafo deste artigo. Após passar pela ferramenta de identificação de similaridade da própria *RAE*, descobri que este artigo tem 9% de similaridade com esses editoriais que foram escritos por mim mesmo e publicados na mesma revista. Ainda que a *RAE* não vá querer processar a si mesma, não vai faltar quem imagine que estou “requeitando” textos antigos quando o que todos querem é ler textos inéditos.

Essa é apenas uma pequena amostra da enorme variedade de situações que ocorrem quando um periódico passa a fazer verificação de similaridade em seus processos editoriais. Além da origem do texto, muitas outras questões também fazem parte da avaliação de similaridade. Uma delas é a quantidade de texto similar. Quanto seria “aceitável”: uma frase, um parágrafo, uma página, meio artigo? Além disso, em qual parte do texto foi encontrada a similaridade? Na introdução, na revisão teórica, na descrição da metodologia ou nas conclusões? A combinação de todas essas possibilidades gera uma boa quantidade de situações que nem sempre têm solução fácil.

Outras questões emergem com a crescente utilização de ferramentas automáticas de detecção de similaridade. Por fazer exclusivamente verificação de similaridade textual, ou seja, identificar semelhanças entre um texto submetido e outros já publicados, essas ferramentas fazem apenas uma checagem automática palavra a palavra, sendo incapazes, por exemplo, de comparar textos publicados em línguas diversas ou mesmo ideias idênticas escritas de modo diferente, coisa que humanos são perfeitamente capazes de fazer com muito mais eficiência.

Um caso digno de relato que ilustra os limites dessas ferramentas aconteceu quando um revisor desconfiou de similaridades em um artigo que havia passado incólume pelo sistema da *RAE*. O revisor pediu nova checagem, e verificou-se que havia identificação em um grande número de citações duplas (algo como: Fulano, 1900; Beltrano, 2000), o que é bastante incomum. O autor escapou da ferramenta ao traduzir *ipsis litteris* trechos de outro artigo, copiando inclusive as referências, mas não conseguiu driblar o olhar atento do revisor. A similaridade era alta, mas em outra língua, e o artigo foi barrado.

Como as ferramentas de detecção de similaridade continuam evoluindo, em breve devem incorporar também recursos de inteligência artificial. Suspeito que o processo vai ficar mais complexo para os periódicos, que terão que se adaptar a uma tipologia ainda mais sofisticada de similaridades.

Em suas diretrizes (Diretrizes *RAE*, s. d.), a *RAE* explica todo o processo de rastreamento de similaridades, indicando claramente que o relatório gerado pela ferramenta é utilizado complementarmente para a tomada de decisão sobre qualquer

artigo. Após esse processo, o artigo é avaliado pela equipe da redação, antes de qualquer comunicação aos autores. Qualquer que seja o resultado obtido no relatório de similaridades, o tom utilizado na comunicação com os autores nunca deve ser de condenação prévia. É sempre melhor fazer pedidos de esclarecimentos sobre os pontos em que a similaridade ultrapassa limites considerados razoáveis e dar a palavra ao autor.

APRENDIZADO CONTINUADO SOBRE UM TEMA COMPLEXO E EM EVOLUÇÃO

Como está dito e reafirmado desde o início deste curto artigo, lidar com o tema das similaridades não é tarefa simples, pois envolve questões culturais num ambiente tecnológico em constante evolução. Os próprios mecanismos de disseminação do conhecimento científico, do qual fazem parte os periódicos acadêmicos, estão se modificando muito rapidamente, por questões técnicas e econômicas.

Além disso, tradicionalmente a Ciência evolui sobre os pilares da livre circulação do conhecimento, da construção de novos conceitos a partir da base preexistente, e da identificação pelos pares sobre o que é, de fato, novo para uma determinada área. Embora esses princípios sejam simples e óbvios, a sua aplicação na prática é bem mais complicada, devido a fatores econômicos, sociais e até mesmo políticos.

Como nos ensina o professor Maggiolini (2014), mesmo a ética tem sido afetada pela crescente digitalização da sociedade. O caso Bauman, mencionado anteriormente, é emblemático, até porque associado a um pensador que deu grandes contribuições para entendermos nossa sociedade interconectada.

Os autores acadêmicos, pressionados pela onda do produtivismo, usam os recursos digitais como podem para atingir os objetivos de ampliar seu impacto nas suas respectivas comunidades científicas. Organismos como o *Committee on Publication Ethics* (COPE) nos fornecem apoio para entender melhor os limites do que podemos e do que não devemos fazer para atingir esses objetivos, e todos os envolvidos com a atividade acadêmica devem tomar contato com os seus princípios de transparência e melhores práticas para publicações científicas (COPE, 2015).

Para encerrar, o fundamental é evitar o comportamento dos “gatos que copiam” (do inglês *copycat*, ou aquele que copia ideias dos outros e as divulga como sendo suas) em qualquer das suas gradações possíveis. Usemos todos os recursos digitais disponíveis, procurando agir sempre de maneira ética e evitando o produtivismo.

REFERÊNCIAS

- Committee on Publication Ethics. (2015). *Principles of transparency and best practice in scholarly publishing*. Recuperado de http://publicationethics.org/files/Principles_of_Transparency_and_Best_Practice_in_Scholarly_Publishingv2.pdf
- Diniz, E. H. (2013). *Ética e bom senso contra o produtivismo*. *RAE-Revista de Administração de Empresas*, 53(4), 331. doi:10.1590/S0034-75902013000400001
- Diniz, E. H. (2014). *Editorial*. *RAE-Revista de Administração de Empresas*, 54(4), 351. doi:10.1590/S0034-759020140401
- Diniz, E. H. (2015a). *Tradução automática, terceiro estágio da revolução digital dos periódicos*. *RAE-Revista de Administração de Empresas*, 55(2), 119. doi:10.1590/S0034-759020150201
- Diniz, E. H. (2015b, Maio/Junho). *Similaridade e plágio: Novos desafios para a gestão de periódicos científicos*. *Editorial*. *RAE-Revista de Administração de Empresas*, 55(3), 239. doi:10.1590/S0034-759020150301
- Diretrizes RAE. (s.d). *Rastreamento de similaridades*. Recuperado de <http://rae.fgv.br/manual-rae>
- Jump, P. (2014). Zygmunt Bauman rebuffs plagiarism accusation. *Times Higher Education*. Recuperado de <https://www.timeshighereducation.com/>
- Machado, A. M. N., & Bianchetti, L. (2011). *(Des)fetichização do produtivismo acadêmico: Desafios para o trabalhador-pesquisador*. *RAE-Revista de Administração de Empresas*, 51(3), 244-254. doi:10.1590/S0034-75902011000300005
- Maggiolini, P. (2014). *Um aprofundamento para o conceito de ética digital*. *RAE-Revista de Administração de Empresas*, 54(5), 585-591. doi:10.1590/S0034-759020140511
- The Visual Communication Guy. (2014). *Did I plagiarize?: The types and severity of plagiarism violations*. Recuperado de <https://visual.ly/community/infographic/education/did-i-plagiarize-types-and-severity-plagiarism-violations>