

# Ciência aberta, dados abertos e código aberto

por Fabio Kon

A HUMANIDADE POSSUI UMA TRADIÇÃO MILENAR DE COMPARTILHAMENTO DO CONHECIMENTO DE FORMA ABERTA. MAS, AO LADO DELA, OUTRA TRADIÇÃO PARALELA DE SEGREDOS MILITARES, COMERCIAIS OU TECNOLÓGICOS SEMPRE ESTEVE PRESENTE COMO FORMA DE TRAZER VANTAGEM SOCIAL, POLÍTICA OU ESTRATÉGICA PARA OS DETENTORES DO CONHECIMENTO.



Por um lado, desde os filósofos gregos da Antiguidade até os cientistas da Renascença, desde os trovadores medievais até os grandes compositores eruditos do Século XIX, a ideia de que o conhecimento gerado pelos intelectuais deveria estar ao alcance

---

**De uma maneira geral, quando há um nível maior de compartilhamento de ideias e abertura do conhecimento, o avanço da ciência é mais rápido e as sociedades tornam-se mais avançadas, ricas e democráticas.**

---

de todos aqueles que quisessem apreciá-lo esteve sempre presente. Realizar cópias das obras dos grandes mestres da pintura é uma tradição de séculos, que vemos com frequência nos museus europeus. A prática de tocar livremente, ao piano, obras dos grandes compositores, seja em audiências públicas ou privadas, existe desde a invenção deste instrumento. O Século XX começou a implementar a ideia, surgida nos séculos anteriores e disseminada com a Revolução Francesa, de que a educação deveria ser universalizada de forma que todo indivíduo tivesse a oportunidade de explorar ao máximo suas capacidades e ter acesso ao conhecimento

produzido por toda a humanidade.

Por outro lado, mecanismos de controle do conhecimento, proteção da informação e até criptografia existem há alguns séculos. No Século XX, no entanto, houve um movimento na di-

reção do fechamento do conhecimento patrocinado por grandes corporações no sentido de privar parte das pessoas do acesso ao conhecimento como forma de gerar receita financeira. Dessa forma, apenas quem pagasse pelo direito de executar uma peça musical receberia autorização para se sentar ao piano em um concerto público e apenas quem pagasse pelo acesso a um artigo científico teria direito de lê-lo. Uma parte significativa dos avanços científicos do século foram motivados por fins militares,

em que a prática de esconder as descobertas do inimigo sempre foi muito forte.

---

**No Século XX, houve um movimento na direção do fechamento do conhecimento patrocinado por grandes corporações no sentido de privar parte das pessoas do acesso ao conhecimento como forma de gerar receita financeira.**

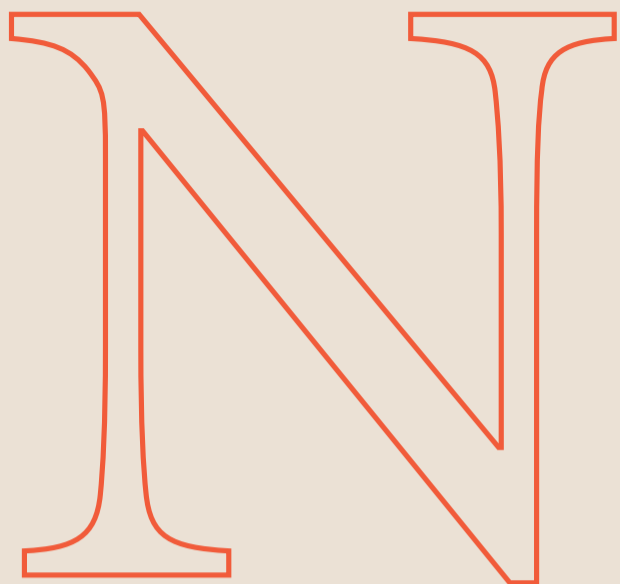
---

**N**

o entanto, o que podemos observar a longo prazo é que, de uma maneira geral, quando há um nível maior de compartilhamento de ideias e abertura do conhecimento, o avanço da ciência é mais rápido e as sociedades tornam-se mais avançadas, ricas e democráticas.

No início do Século XX, os modelos abertos e fechados da prática científica conviviam

lado a lado. Enquanto Santos-Dumont publicou livremente os projetos de seu avião Demoiselle na revista Popular Mechanics em 1910 e conclamou a comunidade a desenvolver, em conjunto, máquinas voadoras cada vez melhores, os irmãos Wright desenvolveram a sua tecnologia em segredo, realizando experimentos longe da atenção do público e da imprensa, e patentearam suas invenções de forma a “proteger sua propriedade intelectual”.



o tocante à disseminação dos resultados das pesquisas científicas, o Século XX testemunhou o surgimento de grandes editoras especializadas na publicação de revistas científicas. A fim de financiar os custos de seleção e editoração de artigos científicos e a distribuição global desses periódicos, passou-se a cobrar quantias significativas por cada artigo individual ou pela assinatura da revista. Bibliotecas acadêmicas de todo o mundo passaram a assinar centenas ou milhares de periódicos. Como não havia concorrência, ou seja, apenas um editor publica uma dada revista, e os cientistas faziam questão de ter acesso a todas as revistas relevantes de sua área, o preço das assinaturas começou a subir de forma significativa. Algumas assinaturas anuais pagas por bibliotecas universitárias hoje em dia chegam a custar dezenas de milhares de dólares.

Chegamos a uma situação na qual (1) o cientista envia o seu trabalho de graça para a revista, (2) o trabalho é avaliado gratuitamente por especialistas voluntários, que escrevem pareceres

---

**Algumas assinaturas anuais pagas por bibliotecas universitárias hoje em dia chegam a custar dezenas de milhares de dólares.**

---

detalhados sem ganhar nada por isso, e (3) todo o trabalho é supervisionado e gerenciado pelos editores-chefe, que também fazem seu trabalho de forma voluntária. Em seguida, todos esses cientistas que trabalharam gratuitamente agora precisam pagar milhares de dólares para ter acesso ao conteúdo publicado nessas mesmas revistas. Os que têm sorte de trabalhar em instituições ricas que possuem assinaturas das revistas em suas bibliotecas ou têm a felicidade de ter

vultosos financiamentos governamentais, como faz a CAPES no Brasil com o seu portal de periódicos, possuem acesso à maior parte desse conhecimento. Já os cientistas de países em desenvolvimento sem um grande financiamento público à ciência, ou mesmo cientistas de instituições periféricas em países desenvolvidos, ficam privados desse acesso à informação.

Felizmente, essa situação *sui generis* começou a ser questionada pela comunidade científica nos últimos anos, e temos presenciado o surgimento de um forte movimento pelo acesso aberto ao conhecimento científico. Revistas científicas de acesso aberto

---

**As principais editoras científicas atualmente permitem ao autor do artigo pagar uma quantia específica a fim de que seu texto seja disponibilizado na Web de forma livre mesmo em uma revista fechada.**

---

passaram a surgir e até mesmo as grandes editoras tradicionais começaram a perceber que precisariam iniciar um movimento nessa direção para se manter no negócio.

As principais editoras científicas atualmente permitem ao autor do artigo pagar uma quantia específica a fim de que seu texto seja disponibilizado na Web de forma livre mesmo em uma revista fechada. Já outras revistas adotam um modelo 100% aberto onde todos os artigos são disponibilizados na Web e os custos de produção da revista são absorvidos pelos autores dos artigos, normalmente com financiamento oriundo

de seus projetos de pesquisa.

Springer Journal of Internet Services and Applications, por exemplo, uma revista criada por iniciativa do CGI.br, SBC e LARC, passou de um modelo fechado baseado em assinatura, em seus três primeiros anos de vida, para um modelo 100% aberto a partir de janeiro de 2013. Nessa revista, cujo editor-chefe é este quem vos escreve, ao lado do Prof. Gordon Blair, todos os artigos são disseminados na Web com uma licença Creative Commons que permite o uso e distribuição irrestrita e uma taxa de publicação é paga pelos autores quando o artigo é aceito. Um financiamento do CGI.br e a política da Springer permitem que autores sem recursos de projetos de pesquisa ou financiamento governamental possam também publicar seus artigos de forma



gratuita. O objetivo da revista é que todo artigo de boa qualidade seja publicado e acessível a qualquer um, em todo o planeta e de forma gratuita, independentemente da capacidade financeira do grupo de pesquisa dos autores.

**Acredito que a tendência, a longo prazo, é o desaparecimento das revistas fechadas no modelo de assinaturas de alto custo e a predominância das revistas de acesso completamente aberto.**

Acredito que a tendência, a longo prazo, é o desaparecimento das revistas fechadas no modelo de assinaturas de alto custo e a predominância das revistas de acesso completamente aberto. Algumas novas publicações, como o Journal of Software Engineering Research and Development (JSERD), recém-lançado pela Comissão Especial de Engenharia de Software da SBC, já estão nascendo no modelo aberto. Vamos torcer para que esse se torne o padrão daqui para frente.

Mas Ciência Aberta não se limita à forma das publicações abertas; precisamos também lutar pelo conteúdo aberto. O que é isso? Um dos pilares da Ciência Experimental é a sua reprodutibilidade.

E ela só se torna reprodutível se os dados e ferramentas utilizados nos experimentos, simulações e análises forem também disponibilizados de forma aberta e livre. Portanto, precisamos de Dados Abertos e Código Aberto, ou seja, Software Livre.



fundamental que os grupos que desejem fazer Ciência de alta qualidade divulguem de forma aberta os dados que utilizaram para realizar os seus experimentos. Open Data não significa somente disponibilizar centenas de megabytes em um site web qualquer, mas também oferecer metadados e documentação de boa qualidade utilizando, sempre que possível, padrões internacionais de codificação de informação, facilitando a reutilização dos dados por terceiros.

Finalmente, as ferramentas de processamento desses da-

dos, que são utilizadas para filtragem, consolidação, análise e visualização da informação, devem ser disponibilizadas como software livre de código aberto, ou seja, um software com código-fonte distribuído segundo uma licença aprovada pela Open Source Initiative (OSI) ou Free Software Foundation (FSF), de preferência uma das licenças populares listadas em [www.opensource.org/licenses](http://www.opensource.org/licenses).

A ideia de que os resultados da pesquisa financiada com dinheiro público devem ser disseminados amplamente de forma

**A ideia de que os resultados da pesquisa financiada com dinheiro público devem ser disseminados amplamente de forma aberta está ganhando cada vez mais força na comunidade científica internacional.**

aberta está ganhando cada vez mais força na comunidade científica internacional. Algumas das principais universidades do planeta, tais como MIT e Harvard, governos, tais como os dos EUA e Inglaterra e agências de fomento, tais como FAPESP, European Commission e NSF, têm começado a incentivar explicitamente a Open Science. No entanto, em algumas áreas ainda existe uma forte resistência de grupos de pesquisa acostumados a esconder os seus dados e procedimentos de trabalho. Em alguns casos, o segredo é uma forma de buscar vantagem comercial – o que, dependendo do contexto, não é necessariamente uma má prática –, às vezes, é uma imposição dos financiadores da pesquisa, em outros casos menos nobres é uma forma

de evitar críticas e esconder possíveis fragilidades do trabalho. Finalmente, em muitos casos, as informações deixam de ser disponibilizadas de forma aberta simplesmente por inércia e falta de uma cultura de dados abertos e software livre.

Por isso, os mecanismos de avaliação da produção dos cientistas não podem se limitar apenas a contar o número de artigos publicados em periódicos de bom nível. É preciso considerar a qualidade da pesquisa em si, bem como o nível de transferência

de conhecimento à sociedade que foi efetivamente obtido. Caso contrário, transformaremos nossos cientistas em eficientes máquinas de produção de papers inúteis.

Cabe a todos nós, produtores e consumidores de Ciência, estudantes, professores, pesquisadores e cidadãos, lutarmos pela universalização do acesso ao conhecimento, cobrando dos cientistas e financiadores públicos da Ciência os pilares da Ciência Aberta: publicação de Acesso Aberto, Dados Abertos e Software Livre. ●



FABIO KON | É Professor Titular em Ciência da Computação do IME-USP e membro da Diretoria da Open Source Initiative.

Agradeço ao pessoal da lista [csl-comunidade@ime.usp.br](mailto:csl-comunidade@ime.usp.br) pelos excelentes comentários e sugestões a uma versão preliminar deste artigo.

