

Preprints na comunicação científica da Física de Altas Energias: análise das submissões no repositório arXiv (2010-2015)

Gonzalo Rubén Alvarez

Doutorando do Programa de Pós-Graduação em Comunicação e Informação da Universidade Federal do Rio Grande do Sul (UFRGS). Mestre em Ciência da Comunicação e Informação pelo PPGCOM/UFRGS.

Sônia Elisa Caregnato

Doutora em Information Studies pela University of Sheffield, Inglaterra. Professora Associada do Departamento de Ciências da Informação da Universidade Federal do Rio Grande do Sul (UFRGS).

<http://dx.doi.org/10.1590/1981-5344/2830>

A circulação de preprints na Física de Altas Energias (FAE) remonta a mais de meio século, tendo como objetivos principais acelerar o processo de comunicação científica entre os pares e estimular o acesso livre à literatura especializada da área. O artigo analisa o conjunto de preprints submetidos às diferentes categorias FAE do repositório temático especializado arXiv no período 2010-2015 que foram posteriormente publicados em revistas peer review. Os indicadores bibliométricos demonstram a potencialidade dos preprints como canal precursor de difusão de resultados científicos visto que 70% das submissões foram em seguida absorvidas pelas principais revistas da FAE. Conclui que o êxito alcançado pelas iniciativas Open Access arXiv e INSPIRE-HEP favoreceu o intercâmbio de informações e conhecimentos entre os pesquisadores. O modelo proposto pela FAE pode incentivar cientistas de áreas com características similares a instalarem repositórios e bancos de dados de preprints para suas disciplinas com o intuito de fortalecer a comunicação das descobertas científicas.

Palavras-chave: *Comunicação científica; Preprints; Open Access; Física de Altas Energias.*

Preprints in scholarly communication of High Energy Physics: analysis of submissions to the arXiv repository (2010-2015)

The preprint circulation in High Energy Physics (HEP) goes back over half a century, with the main objective of accelerating the process of scholarly communication among peers and encourage free access to specialized literature. This paper analyzes a subset of preprints submitted to different HEP categories of the arXiv subject repository from 2010 to 2015 which were subsequently published in peer review journals. Bibliometric indicators show the potential of preprints as a forerunner channel for the diffusion of scientific output since 70% of submissions were then absorbed by leading HEP journals. It concludes that the success of the Open Access initiative arXiv and INSPIRE-HEP has improved the exchange of information and knowledge among researchers. The model proposed by HEP can encourage scientists of areas with similar characteristics to setup repositories and preprint databases to their disciplines in order to strengthen the communication of scientific discoveries.

Keywords: *Scientific communication; Preprints; Open Access; High Energy Physics.*

Recebido em 03.06.2016 Aceito em 11.04.2017

1 Introdução

A produção, disseminação e uso da informação são fundamentais para o desenvolvimento das comunidades científicas. Desde séculos passados, o ato de comunicar através de canais informais e formais se tornou um fator necessário para a construção de novo conhecimento. Na atualidade, o periódico *peer review* é o canal de comunicação formal mais utilizado para publicação das descobertas científicas.

Desde a década de 1960, os físicos teóricos e experimentais, ansiosos para divulgarem os resultados científicos para seus pares de uma forma mais rápida do que a publicação convencional, submetem seus relatórios de pesquisa nos repositórios de *preprints*. Anos mais tarde, os avanços tecnológicos e o advento da *Internet* aceleraram o processo de

comunicação científica e tornaram as pesquisas mais visíveis, uma vez que permitiu o acesso remoto de repositórios e bancos de dados informatizados como arXiv e SPIRES (atualmente INSPIRE-HEP).

A adoção por um sistema de comunicação depende muitas vezes das particularidades da área e do tipo de conhecimento que está sendo transmitido (GOLDSCHMIDT-CLERMONT, 2002). À diferença de outros campos, a Física de Altas Energias (FAE) é uma disciplina que está em contínuo desenvolvimento e que exige grandes investimentos de capital para realizar seus programas experimentais. Nesse sentido, a velocidade é um fator importante na comunicação visto que o atraso no processo de transmissão das informações pode causar uma significativa perda econômica e duplicidade dos experimentos. O fio condutor da inovação FAE na comunicação científica é a sua forte cultura *preprints* (AYMAR, 2009).

Dada a preocupação e a necessidade de comunicação ao longo do tempo, o movimento *Open Access* dentre os membros da comunidade da FAE já não é mais novidade. Por tanto, a importância deste estudo reside na possibilidade de compreender o valor dos *preprints* como mecanismo diferenciado e alternativo na comunicação científica da Física de Altas Energias. Os padrões de comunicação aqui relatados podem ter alguma semelhança com os de outros campos com características similares às da FAE: ciência em rápido desenvolvimento, alta complexidade dos experimentos, grande investimento de capital, pequeno número de laboratórios envolvidos, código de ética tradicional amplamente aceito pelos membros da comunidade científica (GOLDSCHMIDT-CLERMONT, 2002). Esta pesquisa bibliométrica analisa o potencial dos *preprints* na comunicação científica da Física de Altas Energias a partir das submissões recebidas pelo repositório arXiv no período de 2010-2015 e posteriormente publicados em revistas *peer review*.

2 Cultura *preprints* na física de altas energias

Em qualquer área do conhecimento, a realização de pesquisas e a comunicação de seus resultados são atividades inseparáveis (MEADOWS, 1999). No caso da FAE, desde a realização dos primeiros experimentos, os cientistas tiveram a necessidade de transmitir os resultados para a sociedade da qual faziam parte. O diálogo é primordial para o desenvolvimento dessa área destacada pelo trabalho em colaboração. Segundo O'Connell (2000), a FAE sempre teve uma forte cultura de comunicação.

A divulgação das pesquisas e os canais formais e informais de comunicação tornaram-se fatores essenciais na troca de ideias e métodos científicos, sendo que a escolha deles varia segundo a área de conhecimento. Para o avanço da ciência, a submissão de artigos em periódicos científicos com avaliação dos pares ou "*peer review*" torna-se fundamental (MUELLER, 2000). Geralmente, a avaliação de instituições de pesquisa e jovens pesquisadores é realizada com base nas publicações em

prestigiosos jornais *peer-review*. O principal papel das revistas da FAE é ser o “guardião dos registros” (AYMAR, 2009). Entretanto, o surgimento do *World Wide Web* e a criação de bancos de dados específicos de pré-publicações em áreas como a FAE modificaram o comportamento dos cientistas, passando-se para uma “cultura de *preprints*” (MEADOWS, 1999; HURD, 2000; KLING; McKIM, 2000; AYMAR, 2009).

Segundo Søndergaard, Andersen e Hjørland (2003, p. 295): “Apesar de muitas vezes passarem por algum processo de revisão, os *preprints* são documentos distribuídos antes da sua publicação formal e talvez antes da conclusão do processo de avaliação por pares.” Na concepção de Aman (2013), os *preprints* são documentos digitais que foram submetidos a um repositório sem uma revisão prévia por pares. Para Goldschmidt-Clermont (2002), os *preprints* são cópias disseminadas de artigos submetidos para publicação em revistas científicas. Segundo a autora, os *preprints* são documentos temporários que compensam o atraso das publicações e estimulam novas abordagens científicas. Eles são distribuídos para os cientistas que participam em um mesmo campo. Assim que o artigo for publicado, o *preprint* pode e deve ser destruído, referenciando-se apenas a publicação (GOLDSCHMIDT-CLERMONT, 2002).

Na década de 1960, os físicos já submetiam seus relatórios impressos a *clearinghouses*, que são instituições envolvidas com a obtenção, armazenamento e disseminação de documentos não publicados de uma área (SØNDERGAARD; ANDERSEN; HJØRLAND, 2003). Hoje, as versões eletrônicas, como os servidores de *preprints* do *Los Alamos National Laboratory* (LANL) e *European Organization for Nuclear Research* (CERN), são mera e exclusivamente extensões desses serviços (KLING; McKIM, 2000). Com relação à comunicação, Mueller (1994) complementa que os *preprints* se popularizaram como um veículo semiformal naquele período e que em comparação com os canais formais, eles ofereciam maior rapidez e flexibilidade, permitindo a inclusão de comentários. Meadows (1999) certifica que nas pré-publicações impressas, a iniciativa competia ao autor, no caso das pré-publicações eletrônicas, ao leitor. Em termos de inovação, o advento do *fax* acelerou a distribuição desses documentos, mas não foi capaz de diminuir a lentidão na comunicação. Todavia, anos mais tarde, a *Internet* se apresentaria como uma opção eficiente para contornar o atraso entre a divulgação de um trabalho e a sua publicação em um periódico (AMAN, 2013).

Hurd (2000) comenta que já em 1945, Vannevar Bush, em seu artigo *As we may think*, teve a visão de um sistema mais eficaz para o armazenamento e recuperação científica de informações por meio de computadores. A ideia de Bush (1945) era a de inspirar os cientistas das diferentes áreas a se envolverem em um esforço após a Segunda Guerra Mundial para organizar e disponibilizar a quantidade cada vez maior de informações científicas e técnicas através da tecnologia. Bush viu aparentemente o potencial dos computadores e dispositivos de armazenamento que podem ser ligados a sistemas inovadores de indexação para melhor servir informações à comunidade científica (HURD,

2000). Os físicos podem ter se inspirado nas ideias de Bush e optado pela criação de servidores de *preprints* para acelerar o processo de comunicação das pesquisas entre os membros da comunidade.

Ainda, Hurd (2000) aponta que a cultura de *preprints* valoriza a rápida disseminação de resultados preliminares, acelera o processo de comunicação entre os pesquisadores, proporciona a participação ativa dos colégios invisíveis¹, permitindo que prevaleçam as colaborações em grande escala, a dispersão geográfica dos grupos de pesquisa e o acesso à informação por parte das outras áreas que participam nos projetos. A velocidade de transformação na comunicação científica dependerá da aceitação do veículo de cada disciplina.

Søndergaard, Andersen e Hjørland (2003) entendem que o crescente uso e impacto dos canais baseados na *Internet* mudou o fluxo da comunicação científica. Na FAE, o desenvolvimento de banco de dados de *preprints* como nova via de conversação teve um papel vital devido ao descontentamento dos cientistas com atrasos de publicação e problemas de distribuição dos periódicos impressos. Essas dificuldades diminuem a velocidade da comunicação, tendo um efeito adverso sobre a ciência (AMAN, 2013). Sena (2000) explica que Bush na sua obra também apontava o tempo perdido entre a redação do trabalho científico e a sua publicação. Como foi comentado anteriormente, a FAE é um campo que está se desenvolvendo rapidamente e que exige para isso grandes investimentos de capital; nesse sentido, a velocidade é um fator importante na comunicação (GOLDSCHMIDT-CLERMONT, 2002). A autora também comenta que os atrasos que ocorrem no processo de transmissão de informações podem causar uma quantidade significativa de experimentos duplicados, sendo necessária uma rígida vigilância por causa dos gastos envolvidos. Bourdieu (2004) aponta a preocupação dos pesquisadores das grandes revistas americanas da Física com relação à demora dos artigos, já que cinco minutos de atraso podem provocar a perda do benefício de vinte anos de pesquisa. Os editores assumiram o compromisso de reduzirem o atraso e manterem a qualidade do conteúdo, buscando o equilíbrio. Os *preprints* são uma técnica que permite compensar o aspecto negativo desse equilíbrio (GOLDSCHMIDT-CLERMONT, 2002). Meadows (1999) aponta que, apesar da falta de controle de qualidade (uma característica invisível), os servidores de *preprints* bem poderiam constituir um periódico eletrônico plenamente desenvolvido.

O alto custo dos experimentos também teve influência no sistema de comunicação da FAE visto que os trabalhos submetidos aos servidores de *preprints* passam por um forte controle de revisão antes de serem publicados nas revistas da área. De tal modo, um relatório de pesquisa pode ser lido e comentado por dezenas de revisores antes da sua divulgação (KLING; MCKIM, 2000). Nas grandes colaborações, a interação e o consenso entre os cientistas são fatores primordiais para a promoção

¹ “[...] grupo de pesquisadores que está, em um dado momento, trabalhando em torno de um mesmo problema ou área de pesquisa e se comunica sobre o andamento das pesquisas.” (MUELLER, 1994, p. 310).

das pesquisas. O'Connell (2000) afirma que as revistas da área tiveram pouca dificuldade para lidar com a existência dos *preprints*. O autor considera que a publicação em uma revista arbitrada é muito importante, sendo que quase 70% dos *preprints* eventualmente são publicados em periódicos e outros 20% aparecem em anais de congresso.

A ampla aceitação dos repositórios de *preprints* como canal de comunicação diferenciado e alternativo supõe pensar que, os membros da comunidade científica da FAE compartilham os mesmos paradigmas, pois, como expressa Kuhn (1997, p. 219), "Um paradigma é aquilo que os membros de uma comunidade partilham e, inversamente, uma comunidade científica consiste em homens que partilham um paradigma." Conforme Hochman (1994), a comunicação entre as várias comunidades, por vezes, é bastante difícil, o que não acontece dentro de um grupo ou comunidade científica que partilha da mesma tradição, crenças, valores, técnicas etc. O autor explica que, para Kuhn, a aceitação de um paradigma pode ser comprovada com o surgimento de jornais, revistas especializadas, fundação de sociedades científicas, currículos de cursos universitários, citações, livros didáticos etc. Estas são algumas das formas de socialização e comunicação entre os membros do grupo. Devido ao consenso observado na FAE, os *preprints* como meio de comunicação podem ser interpretados como um paradigma criado pela área e utilizados como vias principais de acesso para a literatura publicada nas revistas científicas.

3 Iniciativas Open Access em FAE: SPIRES e ARXIV

A cultura *Open Access* na FAE não é notícia e, como foi dito anteriormente, remonta a décadas, quando os cientistas mandavam *preprints* (manuscritos que ainda não tinham sido publicados em revistas) para seus pares. Dessa conduta, as bibliotecas dos centros de pesquisa classificaram e indexaram esses trabalhos, dando origem ao *Stanford Public Information Retrieval System* (SPIRES), o primeiro servidor de pré-publicações e ao repositório arXiv (GENTIL-BECCOT; MELE; BROOKS, 2010).

Originalmente, SPIRES foi desenvolvido no centro de pesquisa *Stanford Linear Accelerator Center* (SLAC) em 1969, o sistema foi concebido como um sistema de gerenciamento de banco de dados de *preprints* da FAE. No início de 1970, foi adotado pela *Stanford University* para ser utilizado por professores, estudantes e funcionários. Paralelamente, o SLAC e o laboratório *Deutsches Electronen Síncrotron* (DESY) começaram a trabalhar juntos para criar um único banco de dados na área. Com o lançamento da *World Wide Web* em 1991, SPIRES tornou-se o primeiro banco de dados na *Web* (O'CONNELL, 2010).

No SPIRES, as citações a *preprints* são agregadas às citações das versões publicadas, tratando as duas como uma entidade única. O estudo de Gentil-Beccot, Mele e Brooks (2010) demonstrou que os artigos da FAE submetidos a um repositório *Open Access*, como o arXiv, e posteriormente

publicados em revistas, recebem cinco vezes mais citações do que os artigos que são publicados somente em revistas ou em repositório e que 20% das citações aos artigos da área, em um período de dois anos, ocorre antes da publicação. Em conjunto, esses resultados levaram a três conclusões gerais sobre a comunicação científica na FAE: adoção *Green Open Access*² para as publicações; maior visibilidade e reconhecimento dos pares devido à livre e imediata circulação de trabalhos; valor dos repositórios de *preprints* como meio de discurso complementar dos periódicos científicos. Atualmente, uma das ferramentas de busca mais utilizada pela área é o INSPIRE. Baseado na experiência SPIRES, o INSPIRE (INvenio + SPIRES) é um sistema de informação da FAE construído pelos laboratórios CERN, DESY, Fermi National Accelerator Laboratory (FERMILAB) e SLAC. Ele combina o sucesso do banco de dados SPIRES com a tecnologia da biblioteca digital Invenio desenvolvido no CERN. INSPIRE é gerido pela colaboração dos quatro laboratórios e interage de perto com outros sistemas como arXiv, NASA-ADS, PDG, HEPDATA. Representa uma evolução natural da comunicação científica, construída sobre sistemas de informação bem-sucedidos baseados em uma comunidade e fornece uma visão para a gestão da informação em outros campos da ciência (INSPIRE-HEP, 2016b, documento *on-line*).

Em agosto de 1991, Paul Ginsparg no LANL de Novo México, Estados Unidos, criou o primeiro servidor de *e-print* (*electronic pre-print*), o arXiv (SENA, 2000; MUELLER, 2006). O arXiv é um repositório temático especializado de *preprints* de acesso aberto. As áreas abrangidas incluem Física, Matemática, Ciências da Computação, Ciências não-lineares, Biologia Quantitativa e Estatística. O servidor é mantido e operado pela biblioteca da *Cornell University*, com orientação do Conselho Consultivo Científico do arXiv. As submissões são analisadas por especialistas da área, verificando-se a relevância das mesmas para a comunidade científica. O arXiv complementa o sistema de publicação tradicional e proporciona a divulgação imediata e gratuita da informação em acesso aberto. O foco do arXiv é a disponibilidade perpétua das submissões. As submissões aceitas recebem um ID de arXiv. Os autores podem fazer atualizações das suas pesquisas a partir da inserção de novas versões. As versões antigas não podem ser removidas (arXiv, 2015, documento *on-line*).

Sena (2000) acrescenta que a essência do modelo de comunicação *Open Access* na FAE é a livre circulação dos manuscritos, sendo que os especialistas podem discutir, criticar e complementar o trabalho dos pares antes da publicação nos importantes periódicos da área. Muitos pesquisadores fazem as atualizações no arXiv das versões revisadas e corrigidas durante o processo de publicação nos periódicos (GENTIL-BECCOT; MELE; BROOKS, 2010). Na visão de O'Connell (2000) e Mueller (2006) o ingrediente chave na criação de bases de dados de *preprints* foi

² O termo indica a disponibilidade gratuita de publicações acadêmicas em repositórios digitais de acesso livre. No caso da FAE, a submissão a esses repositórios, como por exemplo, o arXiv, não é obrigatória por universidades ou agências de financiamento, mas é uma escolha dos autores buscando o reconhecimento dos colegas e uma maior visibilidade (GENTIL-BECCOT; MELE; BROOKS, 2010).

a centralização. Os cientistas da área da física e outras áreas relacionadas, localizados em qualquer parte do mundo, poderiam enviar seus trabalhos para um repositório central e ao mesmo tempo se manter atualizados com relação às últimas pesquisas na FAE. Mueller (2006) acrescenta que embora os trabalhos enviados não tivessem sido avaliados por pares, o sistema verificava alguns pontos de maneira a garantir uma qualidade mínima (por exemplo, filiação do autor). Assim, os cientistas remetiam seus trabalhos para a base de dados do LANL, ao mesmo tempo em que os submetiam para as editoras das principais revistas científicas da área. Posteriormente, a ascensão do serviço de redes de computadores consolidou a prática. Segundo Aymar (2009), três inovações marcaram avanços cruciais na comunicação científica da FAE: o surgimento da base de dados SPIRES, a concepção do repositório especializado arXiv e a invenção da *Web* no CERN. Graças à sua cultura *preprints*, a FAE é uma disciplina que valoriza a rápida disseminação e acesso à informação (AYMAR, 2009). Como canal de comunicação informal, os *preprints* podem ser considerados os antecessores da publicação formal. Sua relação com os periódicos da área está firmemente estabelecida na essência da pesquisa científica.

4 Metodologia

O *corpus* de análise deste estudo bibliométrico é composto pelos *preprints* submetidos ao repositório arXiv no período de 2010-2015 e classificados em uma das quatro categorias FAE: **hep-ex** para *high energy experimental physics*; **hep-lat** para *studies of lattice field theory*; **hep-ph** para *particle phenomenology*; **hep-th** para *string, conformal and field theory*. A coleta das informações foi realizada em maio de 2016 na base de dados INSPIRE-HEP. Esse sistema de acesso aberto para o campo da Física de Altas Energias foi escolhido uma vez que tem cobertura completa dos *preprints* submetidos em arXiv, incluindo informações sobre publicação. Na estratégia de busca foi utilizado o campo 037 do código MARC da base que contém dados referentes ao número de relatório e categoria do *preprint* submetido ao arXiv. O uso do operador booleano asterisco (*) possibilitou a recuperação dos *preprints* com qualquer número de relatório arXiv que contenha, por exemplo, a raiz arXiv:10, o que significa que a submissão foi realizada em arXiv em 2010 (uma vez que a primeira parte do número arXiv refere-se ao ano). A amostra de *preprints* submetidos às diferentes categorias FAE do arXiv no período de 2010-2015 foi obtida a partir da seguinte expressão de busca: (037:arXiv:10* or 037:arXiv:11* or 037:arXiv:12* or 037:arXiv:13* or 037:arXiv:14* or 037:arXiv:15*) and (037:hep-ex or 037:hep-lat or 037:hep-ph or 037:hep-th). Dados correspondendo aos *preprints* submetidos em 2015 e subsequentemente publicados em revistas arbitradas foram utilizados para analisar a concentração dos artigos por veículo de comunicação. A partir da opção *Citesummary* foi possível conhecer a quantidade de *preprints* publicados em periódicos *peer-review*.

Por meio das referências dos documentos geradas pela ferramenta de formatação *BibTex* da base INSPIRE-HEP, realizou-se a coleta dos dados sobre os veículos de publicação dos 6.415 *preprints* de 2015. Indicadores de produção científica foram gerados através da análise das variáveis Ano de submissão (*Year*), Categoria do *preprint* em arXiv (*primaryClass*) e Periódico de publicação do *preprint* (*Journal*). Com relação ao tratamento dos dados, foi utilizado o programa Microsoft Excel 2007 para a confecção de tabelas e gráficos.

5 Resultados

Com base no conjunto de *preprints* submetidos às diferentes categorias da FAE no período de 2010-2015 e posteriormente publicados em periódicos *peer review*, os resultados apresentados na Tabela 1 manifestam a potencialidade desse tipo de literatura cinzenta na FAE visto que 70% das submissões foram absorvidas mais tarde pelas principais revistas da área, coincidindo com os dados da pesquisa de O'Connell (2000). Em nível individual, as categorias *hep-th* e *hep-ph* apresentaram os percentuais de *preprints* publicados em periódicos *peer review* mais altos, 81% e 69% respectivamente.

Tabela 1 – *Preprints* submetidos no arXiv no período 2010-2015 e posteriormente publicados em revistas *peer review*

Ano	hep-ex			hep-lat			hep-ph			hep-th			Total		
	Ns	Np	%	Ns	Np	%	Ns	Np	%	Ns	Np	%	Ns	Np	%
2015	849	479	56	624	259	42	4.939	3.098	63	3.714	2.579	69	10.126	6.415	63
2014	923	507	55	589	259	44	4.630	3.396	73	3.618	3.028	84	9.760	7.190	74
2013	924	452	49	719	283	39	4.538	3.210	71	3.318	2.811	85	9.499	6.756	71
2012	1.084	616	57	579	305	53	4.541	3.240	71	3.298	2.745	83	9.502	6.906	73
2011	1.110	499	45	589	261	44	4.620	3.210	69	3.332	2.743	82	9.651	6.713	70
2010	696	309	44	573	250	44	4.205	2.813	67	3.366	2.764	82	8.840	6.136	69

Fonte: INSPIRE-HEP (2016a).

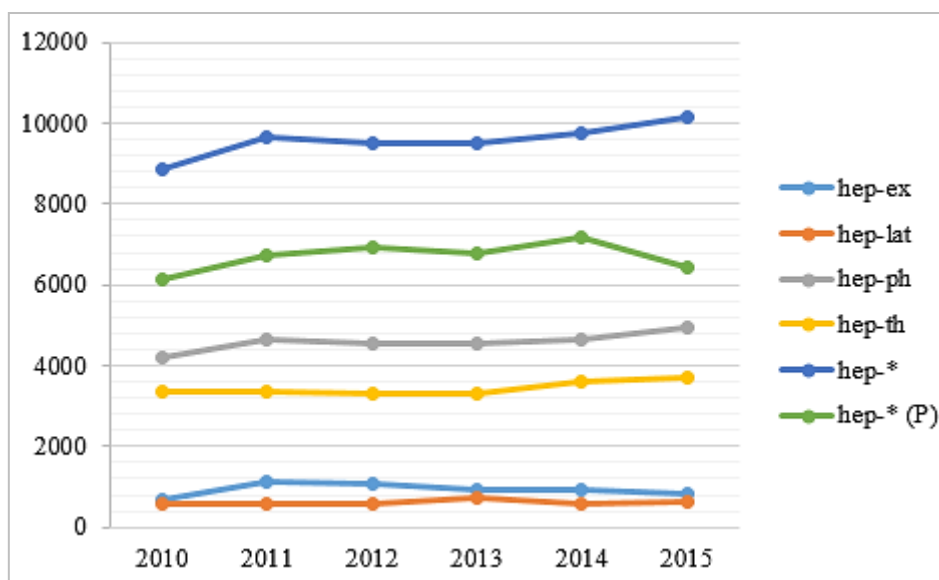
Legenda: **hep-ex** para high energy experimental physics; **hep-lat** para studies of lattice field theory; **hep-ph** para particle phenomenology; **hep-th** para string, conformal and field theory; **Ns** Número de preprints submetidos a arXiv; **Np** Número de preprints posteriormente publicados em periódicos peer-review; **%** Percentual de preprints publicados em periódicos peer-review. Dados sobre os preprints de 2015 posteriormente publicados foram utilizados para analisar a dispersão dos artigos por título de periódico.

Com relação à evolução temporal dos *preprints* submetidos às diferentes categorias da FAE no período de 2010-2015 e posteriormente publicados em periódicos *peer-review*, infere-se que a regularidade do número de submissões anuais representada pela baixa variação percentual manifesta uma consolidada adoção do sistema como canal precursor de comunicação das descobertas dentre os membros da comunidade científica da FAE, confirmando o comportamento observado por Mele *et al.* (2006) ao analisar a evolução histórica do número de

submissões em SPIRES entre 1991 e 2005. Esse indicador retrata a coexistência pacífica de dois canais de difusão com características diferentes, os *preprints* e os artigos de revista.

Na Figura 1, aprecia-se o desempenho das submissões da FAE em arXiv por categoria (hep-ex, hep-lat, hep-ph e hep-th) assim como o número total (hep-*) durante o período estudado. Os *preprints* posteriormente publicados em revistas *peer-review* são indicados com 'P'. Com relação à leve queda destes últimos registrada em 2015 pode-se argumentar que alguns *preprints* submetidos no final de dito ano em arXiv podem se encontrar ainda no processo editorial de revistas *peer-review*, coincidindo com os resultados de Mele et al. (2006).

Figura 1 – Dinâmica de crescimento de *preprints* submetidos no arXiv no período 2010-2015 e posteriormente publicados em periódicos *peer review*



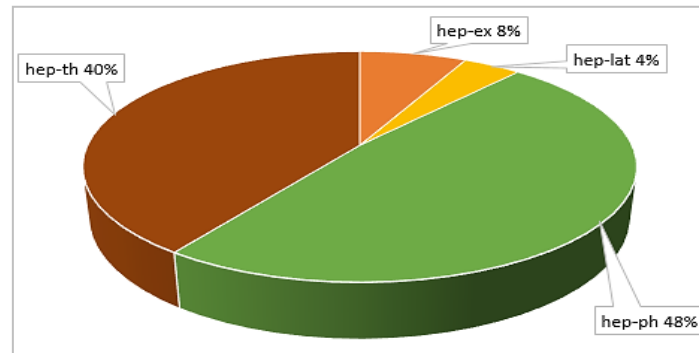
Fonte: INSPIRE-HEP (2016a).

Legenda: **hep-ex** *Preprints high energy experimental physics*; **hep-lat** *Preprints studies of lattice field theory*; **hep-ph** *Preprints particle phenomenology*; **hep-th** *Preprints string, conformal and field theory*; **hep-*** *Total de preprints das quatro categorias FAE submetidos no arXiv*; **hep-* (P)** *Total de preprints das quatro categorias FAE posteriormente publicados em periódicos peer review*.

A Figura 2 apresenta a distribuição por categoria FAE (hep-ex, hep-lat, hep-ph e hep-th) dos 6.415 *preprints* de 2015 posteriormente publicados em periódicos *peer review* nos quais a pesquisa foi baseada. O número de artigos em cada categoria *hep* é proporcional ao tamanho da subárea da FAE que ela representa. A predominância das publicações das categorias *hep-ph* (48%) e *hep-th* (40%) pode ser justificada pelo fato delas englobarem uma maior quantidade de assuntos de diferentes campos da Física de Altas Energias. Os artigos relacionados às pesquisas experimentais (hep-ex), envolvendo grandes colaborações internacionais,

representam apenas 8% do total de publicações. A categoria *hep-lat* (4%) é a menos representativa dentre as subáreas da FAE dada a insuficiência de grupos no mundo com suporte computacional de alto desempenho suficiente para a realização dos estudos.

Figura 2 – Distribuição por categoria FAE dos *preprints* de 2015 posteriormente publicados em revistas *peer-review*

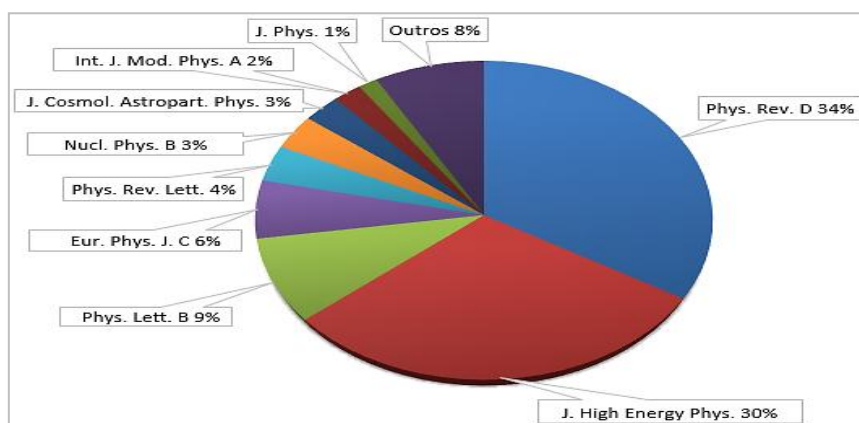


Fonte: INSPIRE-HEP (2016a)

Legenda: **hep-ex** para *high energy experimental physics*; **hep-lat** para *studies of lattice field theory*; **hep-ph** para *particle phenomenology*, **hep-th** para *string, conformal and field theory*.

Outras iniciativas *Open Access* na FAE como o *Sponsoring Consortium for Open Access Publishing in Particle Physics* (SCOAP³) pretendem converter a literatura publicada nos principais periódicos da área para acesso aberto, redirecionando as taxas das assinaturas das revistas para pagamento dos avaliadores *peer-review*. Algumas editoras como a Elsevier, Springer e IOP Publishing já estão participando do SCOAP³.

Figura 3 – Distribuição dos *preprints* de 2015 posteriormente publicados por título de periódico favorecido



Fonte: INSPIRE-HEP (2016a).

A Figura 3 apresenta as revistas que receberam *preprints* de 2015 submetidos em alguma das quatro categorias FAE no arXiv para

publicação. Verifica-se que quase 90% dos 6.415 artigos são publicados em apenas sete periódicos de quatro editoras, sendo eles Physical Review D e Physical Review Letters (APS), Journal of High Energy Physics e European Journal Physical C (Springer), Nuclear Physics B e Physics Letters B (Elsevier), Journal of Cosmology and Astroparticle Physics (IOP Publishing). Percebe-se como a alta concentração de publicações em poucas revistas facilita a iniciativa de acesso aberto proposta pelo modelo SCOAP³.

O fato de os autores optarem por submeter seus trabalhos em repositórios *Open Access* por causa do maior volume de citações recebidas conforme observado na pesquisa de Gentil-Beccot, Mele e Brooks (2010), pode estar influenciando as editoras a flexibilizar seu controle sobre o acesso aos textos que publicam dado que o maior número de citações eleva o Fator de Impacto (FI) e aumenta o prestígio e as vendas dos periódicos (MULLER, 2006).

6 Considerações finais

A disponibilização dos resultados de pesquisas ainda não publicadas em periódicos pelos físicos remonta à década de 1960 e tem como objetivos principais acelerar o processo de comunicação entre os pares e estimular o acesso livre à literatura científica. Anos mais tarde, o surgimento da Internet, a criação de repósitorios e banco de dados informatizados e o atraso das publicações nas revistas especializadas terminaram por confirmar a “cultura *preprints*” na FAE. Essa prática de comunicação proporciona aos pesquisadores da área maior rapidez e flexibilidade no diálogo e estimula a inserção de comentários, tornando o desenvolvimento científico mais consensual e participativo. Como canal de comunicação, os *preprints* podem ser considerados os antecessores da publicação formal. Sua relação com os periódicos está firmemente estabelecida na essência da pesquisa científica, conforme observado neste estudo visto que 70% dos *preprints* submetidos em arXiv no período de 2010-2015 foram posteriormente publicados em revistas arbitradas da área. Em termos de inovação, o êxito alcançado pelas iniciativas *Open Access* arXiv e SPIRES (INSPIRE) na Física de Altas Energias pode ter estimulado cientistas de outros campos científicos a instalarem repositórios e bancos de dados de *preprints* para suas disciplinas com o intuito de melhorar a comunicação das descobertas. É o caso do *Cognitive Sciences Eprint Archive* (CogPrints) que indexa *e-prints* de Psicologia, Lingüística, Neurociências e de algumas áreas da Ciência da Computação, Filosofia e Biologia e do *Research Papers in Economics* (RePEc) de Economia. Por último, a natureza das investigações, o tipo de conhecimento a ser transmitido e a necessidade dos cientistas têm influência na escolha dos canais de comunicação dos resultados de pesquisa e, conseqüentemente, na legitimação das descobertas através da avaliação científica.

Referências

- AMAN, V. The potential of preprints to accelerate scholarly communication: a bibliometric analysis based on selected journals. *arXiv preprint arXiv:1306.4856*, 2013. Disponível em: <<http://arxiv.org/abs/1306.4856>>. Acesso em: 16 abr. 2015.
- ArXiv. *About ArXiv*. 2015. Disponível em: <<http://arxiv.org/>>. Acesso em: 22 maio 2015.
- AYMAR, R. Scholarly communication in high-energy physics: past, present and future innovations. *European Review*, v. 17, n. 1, p. 33-51, 2009.
- BOURDIEU, P. *Os usos sociais da ciência: por uma sociologia clínica do campo científico*. São Paulo: Editora UNESP, 2004.
- BUSH, V. As we may think. *Atlantic Monthly*, v. 176, n. 1, p. 101-108, 1945. Disponível em: <<http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>>. Acesso em: 12 jun. 2015.
- GENTIL-BECCOT, A.; MELE, S.; BROOKS, T. C. Citing and reading behaviours in high-energy physics. *Scientometrics*, v. 84, n. 2, p. 345-355, 2010.
- GOLDSCHMIDT-CLERMONT, L. Communication patterns in high-energy physics. *High Energy Physics Libraries Webzine*, n. 6, 2002. Disponível em: <http://scholar.google.com.br/scholar?cluster=14704071346888558677&hl=pt-BR&as_sdt=0,5>. Acesso em: 23 maio 2015.
- HOCHMAN, G. A ciência entre a comunidade e o mercado: leituras de Kuhn, Bourdieu, Latour e Knorr-Cetina. In: PORTOCARRERO, V. (Org.). *Filosofia, história e sociologia das ciências: abordagens contemporâneas*. Rio de Janeiro: Fiocruz, 1994.
- HURD, J. M. The transformation of scientific communication: a model for 2020. *Journal of the American Society for Information Science*, v. 51, n. 14, p. 1279-1283, 2000.
- INSPIRE-HEP. *HEP Search: High-Energy Physics Literature Database*. 2016a. Disponível em: <<https://inspirehep.net/>>. Acesso em: 13 maio 2016.
- INSPIRE-HEP. *About INSPIRE*. 2016b. Disponível em: <<http://inspirehep.net/info/general/project/index>>. Acesso em: 13 maio 2016.
- KLING, R.; McKIM, G. Not just a matter of time: field differences and the shaping of electronic media in supporting scientific communication. *Journal of the American Society for Information Science*, v. 51, n. 14, p. 1306-1320, 2000.
- KUHN, T. S. *A estrutura das revoluções científicas*. São Paulo: Perspectiva, 1997.
- MEADOWS, J. *A comunicação científica*. Brasília, DF: Briquet de Lemos, 1999.

MELE, S. *et al.* Quantitative analysis of the publishing landscape in High-Energy Physics. *Journal of High Energy Physics*, v. 12, p. 1-25, 2006.

MUELLER, S. P. M. O impacto das tecnologias de informação na geração do artigo científico: tópicos para estudo. *Ciência da Informação*, Brasília, v. 23, n. 3, p. 309-317, set./dez. 1994.

MUELLER, S. P. M. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. (Org.). *Fontes de informação para pesquisadores e profissionais*. Belo Horizonte: Editora UFMG, 2000. p. 20-34.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência da Informação*, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006.

O'CONNELL, H. B. *The horse that drank*: electronic communication and the High Energy Physics community. SLAC-PUB-8566, 2000. Disponível em: <<http://slac.stanford.edu/cgi-wrap/getdoc/slac-pub-8566.pdf>>. Acesso em: 15 abr. 2015.

O'CONNELL, H. B. A Tony Thomas-inspired guide to INSPIRE. In: MELNITCHOUK, W. (Ed.). *American Institute of Physics Conference Series*. Batavia: AIP, 2010. p. 265-270. Disponível em: <<http://lss.fnal.gov/archive/2010/conf/fermilab-conf-10-061-bss.pdf>>. Acesso em: 25 abr. 2015.

SENA, N. K. Open archives: caminho alternativo para a comunicação científica. *Ciência da Informação*, Brasília, v. 29, n. 3, p. 71-78, set./dez. 2000.

SØNDERGAARD, T. F.; ANDERSEN, J.; HJØRLAND, B. Documents and the communication of scientific and scholarly information: revising and updating the UNISIST model. *Journal of Documentation*, v. 59, n. 3, p. 278-320, 2003.