

Big Data, informação aberta para transferência de tecnologia

Big Data, open information for technology transfer

Ana Ponte

M.Sc. Economia e Gestão de Ciência, Tecnologia e Inovação (ISEG/UTL);
Unidade de Parcerias e Cooperação – Agência Nacional de Inovação (ANI) –
Lisboa, Portugal.
aponte@ani.pt

Resumo

A gestão da informação está atualmente sujeita a restrições legais, sobretudo a informação que diz respeito ao conhecimento gerado e potenciador de vantagem económica, como a registada em patentes. Tais restrições legais integram-se num conceito de ciclo da inovação clássico, no qual assentam os parâmetros atuais de financiamento científico e de exploração comercial dos resultados de desenvolvimento tecnológico e de inovação. A inovação aberta, como novo modelo de gestão da inovação, tem desafios inerentes – de ordem legal, cultural e organizacional - ao modelo clássico do ciclo de inovação.

A sessão descrita neste artigo expõe casos concretos de implementação da inovação aberta sobre a utilização de grandes quantidades de dados associados a patentes. Ilustra-se, pois, com base nas verificações desta sessão, os benefícios que a inovação aberta com base em dados de patentes podem trazer ao desenvolvimento sustentável, passíveis de representarem mais-valias sociais e económicas avultadas, e que deverão justificar uma abordagem proativa na resolução dos desafios em causa, nomeadamente a nível da definição de estratégias de inovação aberta.

Da mesma forma, as ferramentas de tratamento de *big data* já em desenvolvimento, e aqui também ilustradas, oferecem garantias à validação desta forma de inovação na resolução de problemas sociais concretos.

Palavras Chave:

Inovação aberta, inovação frugal, *big data*, patentes, desenvolvimento sustentável.

Abstract

The management of information is currently subject to legal restrictions, particularly information relating to generated knowledge enhancing economic advantage, as the one registered in patents. Such legal restrictions derive from a classical concept of the innovation cycle, which provides a basis for the current parameters of Science funding and commercial exploitation of technological development and innovation results. Open innovation, as a new model of innovation management, brings along inherent challenges – at legal, cultural and organizational levels - to the classical model of the innovation cycle.

This paper describes a discussion session that has disclosed specific cases of open innovation implementation, based on the use of big data extracted from patents. Upon the findings of that working session, it intends to showcase the benefits that open innovation based on patent data can bring to sustainable development, likely to represent large social and economic gains, and which may, therefore, justify a proactive approach to the resolution of the mentioned challenges, in particular as regards the drafting of open innovation strategies.

Alongside, big data processing tools already in development, as also illustrated herein, provide assurance as to the validation of this form of innovation to solve specific social problems.

Key Words:

Open innovation, frugal innovation, big data, patents, sustainable development.

Introdução

A inovação aberta, um campo em exploração ao nível interno das organizações empresariais, mas ainda em fase de discussão pública em meios institucionais mais abrangentes, impõe desafios – legais, operacionais, culturais – ao ciclo da inovação clássica. Os aspetos de propriedade intelectual constituem um dos maiores desafios inerentes ao tema, porquanto a utilização de ideias e tecnologias externas no processo de inovação está sujeita, tradicionalmente, a restrições, resolvidas com recurso a soluções que vão desde a autorização à cobrança de *royalties* ou a alianças corporativas.

No entanto, em atuais contextos mais favoráveis à adoção deste modelo (e.g. países como a Índia ou o Brasil), a inovação aberta tem-se provado geradora de um conjunto de benefícios sociais e económicos que justificariam uma abordagem proativa e empenhada na resolução dos desafios em causa. Casos concretos permitem já avaliar um significativo contributo para o desenvolvimento sustentável de comunidades alargadas, especialmente em meios mais carenciados. O seminário “*Big data, desenvolvimento sustentável e a ciência aberta*”, realizado no dia 23 de junho de 2016, em Lisboa, sob organização do Instituto Francês e do Instituto de Higiene e Medicina Tropical da Universidade Nova de Lisboa e com o apoio da Fundação brasileira Fiocruz e da Agência Nacional da Inovação de Portugal, teve como objetivo apresentar casos concretos capazes de evidenciar a relevância de que a inovação aberta pode revestir-se em diferentes contextos. Adicionalmente, o evento mostrou também os meios já disponíveis, e em desenvolvimento, para promover a inovação aberta com base em grandes quantidades de dados e informações, tais como aqueles que existem hoje em dia nas patentes registadas em vários repositórios pelo mundo.

Assim, a sessão *Big data, informação aberta para transferência de tecnologia* deste evento incidiu sobre a apresentação de dois casos práticos de aplicação da inovação aberta sobre dados de patentes e sobre a exposição de duas vertentes de aplicativos informáticos especificamente desenvolvidos para o tratamento destas grandes quantidades de dados. Manuel Durand-Barthez, Conservador Geral da *Unité Régionale de Formation à l'Information Scientifique et Technique* (URFIST) de Paris, apresentou um caso de inovação frugal com base em informação aberta; Sérgio Matos, Investigador Associado do Instituto de Eletrónica e Engenharia Informática da Universidade de Aveiro, abordou a mineração de patentes para descoberta de conhecimento sobre fármacos; Abdelkader Baaziz, Professor Associado da *École Nationale Supérieure e Management d'Alger*, Universidade de Tipaza, apresentou um outro caso concreto de transferência de tecnologia com informação aberta, desta feita na área da exploração petrolífera; e, por fim, a intervenção de David Reymond, Leitor da área de Ciências da Informação

e Comunicação da Universidade de Toulon, incidiu sobre o mapeamento de *big data*.

Descrevem-se abaixo os pontos principais destas apresentações, pretendendo-se destacar a verificação do potencial da inovação aberta para a dinamização dos sistemas nacionais de inovação e para o desenvolvimento sustentável global, tal como transmitido pelos oradores e que constituiu também o objetivo da sessão em causa.

Inovação frugal com base em informação de patentes: casos de sucesso

Na apresentação de Manuel Durand-Barthez, expondo um caso de inovação frugal com informação aberta, esteve patente o forte potencial da inovação frugal para uma outra forma de inovação: a inovação social. Com efeito, Durand-Barthez começou por apontar que os principais canais de informação científica – tais como artigos de investigação ou teses – não são os vetores preferenciais da informação técnica para resolução de problemas, nomeadamente aquela consultada por engenheiros, dado serem de difícil e, muitas vezes, dispendioso acesso. Ao contrário, a patente, um documento de carácter essencialmente técnico, muito concreto, sintético e de livre acesso, constitui uma fonte importante de informação aberta para novas aplicações a problemas sociais concretos.

Porém, o registo de patentes nos países desenvolvidos tende a inserir-se num sistema *top-down* na perspectiva da sua venda aos países menos desenvolvidos. No entanto, a abertura da Propriedade Intelectual e Industrial promovida já em alguns países, tais como a Índia, a China ou o Brasil, permite criar um sistema simétrico, no sentido em que parte de problemas existentes entre os consumidores para chegar, através da consulta dessa informação, a possíveis soluções que possam ser adaptadas, criando um sistema de “inovação reversa” particularmente útil às classes sociais desfavorecidas. Exemplos de aplicação desta forma de inovação, frugal e quase imediata, capaz de resolver problemas medianamente complexos, são métodos de deteção, diagnóstico e luta contra doenças (paludismo, tuberculose, filariose e outras), aplicações de vigilância epidemiológica ou de proteção ambiental.

Naturalmente, não pode ser negligenciado o princípio fundamental do Retorno sobre o Investimento (ROI) associado às patentes, constituindo a contrapartida financeira para a publicação da descrição da inovação. Porém, importa salientar que o registo da patente tem uma incidência geográfica inerente e que, por outro lado, o conceito incluído na patente é protegido contra uma utilização em escala industrial.

Deste modo, a vantagem desta forma de inovação frugal centra-se, não numa atividade inventiva de alta intensidade,

mas na informação e inspiração de indivíduos para a adaptação de inovações existentes e para a conceção de processos simples, num ambiente que não o de negócios ou empresarial, mas sim em contextos de uso sem fins lucrativos, nomeadamente os que servem propósitos humanitários, pelo que forçosamente desenvolvidos sem recurso a meios onerosos ou excessivamente sofisticados.

Particularmente favorável a esta forma de inovação frugal é a proteção das invenções por modelo de utilidade, ao invés da alternativa de proteção por patente.

Tal inscreve-se, conforme concluído por Durand-Barthez, num quadro dinâmico de ciência aberta em favor do progresso social, que não promove a anterioridade clássica, implicando, pois, uma perspetiva positiva, não protecionista.

Um outro caso de inovação reversa, aplicada à engenharia, foi apresentado por Abdelkader Baaziz. Uma vez mais com base na extração da informação de patentes, este caso ilustra a geração de oportunidades de transferência de tecnologia com vista à resolução de problemas concretos, desta feita pelo lado da Indústria, nomeadamente ao nível da otimização do desempenho de uma broca de exploração petrolífera.

Dependendo o sucesso da perfuração para exploração petrolífera de fatores como o tipo de solo, a pressão e a velocidade de perfuração e de sucção do petróleo ou a profundidade do lençol de hidrocarbonetos, e devendo ser tidos em conta aspetos financeiros associados, quer à seleção das brocas e outros dispositivos, quer aos riscos da própria operação, pode tornar-se demasiado oneroso solucionar determinado problema encontrado na intervenção por recurso a alterações significativas, quer geográficas, quer de meios. Neste sentido, a consulta de informações de patentes relacionadas com a fabricação e otimização de instrumentos permite, desde logo, identificar os líderes e as tendências de mercado na área, nomeadamente para possível estabelecimento de parcerias (*joint ventures*) com base em transferência de tecnologia e uso legal da patente, e mesmo constituir um inventário de tecnologias com registo das datas da sua utilização livre. Aplicativos informáticos como o *MatheoSoftware* ou o *Patent2Net* (desenvolvido por um outro orador do painel, David Reymond) são utilitários convenientes para estes fins, apesar de apenas o segundo ser de acesso aberto.

O caso concreto trazido por Baaziz para ilustrar a utilidade de complementar informação de patentes com técnicas de engenharia reversa para a resolução de problemas centrou-se na necessidade de uma ferramenta de perfuração eficaz e otimizada para aplicação numa formação geológica particular. Para o efeito, a pesquisa de patentes relacionadas com o problema em questão resultou na análise de 7259 patentes. Quando organizadas por classes, a maioria destas patentes referia-se ao fabrico de brocas e outros componentes associados à perfuração, à composição dos compo-

entes cortadores em particular (designadamente às inserções de diamante natural ou de compacto policristalino de diamante - PDC), à otimização das técnicas de corte de elementos das brocas (inserções, formatos, etc.) e a processos de otimização do desempenho das brocas.

Uma outra análise, por data de publicação das patentes, permite identificar as que se referem a tecnologias obsoletas, as que se referem a tecnologias livres (por já ter caducado a proteção da patente) e aquelas que se referem a tecnologias ainda protegidas. Apenas no caso destas últimas se justificará o recurso à identificação das empresas que submeteram a patente a fim de estabelecimento de eventuais parcerias para utilização da tecnologia; nos outros dois casos, há margem para oportunidades de transferência de tecnologia por simples recurso à informação das patentes. No entanto, um aspeto a ter em conta na busca pela utilização destas informações é a eventual criação de lacunas jurídicas na gestão das patentes em situações de fusão e aquisição entre entidades empresariais. Tais situações podem fazer cair as patentes no domínio público antes de terminado o tempo de proteção inicialmente previsto, nomeadamente se houver falta de pagamento das taxas anuais por extinção da entidade responsável. É de salientar que estas situações são explicitamente indicadas em plataformas como a “Espacenet”, o que contribui para a não estagnação dos mercados tecnológicos face a situações como estas, possibilitando a utilização legal do conteúdo destas patentes sem pagamento de *royalties*.

No caso concreto apresentado, foi possível, por recurso a informação das patentes analisadas e por técnicas de engenharia reversa, adaptar e reparar 3 vezes uma broca de PDC para a intervenção em causa. Tal representou uma poupança de 40% do custo face a uma nova broca; por outro lado, a taxa de penetração foi significativamente aumentada (de 16,8 m/h para 22,9 m/h) em condições de segurança utilizando a referida broca otimizada, o que permitiu um ganho de tempo de mais de três dias de perfuração, com poupança de cerca de 3 milhões de dólares. Obteve-se ainda economias de custo para o cliente na ordem das centenas de milhares de dólares.

Desta forma, o processo de análise levado a cabo, conjugando informações de patentes com processos cognitivos associados à técnica da engenharia reversa, permitiu explorar novos caminhos para a inovação e a transferência de tecnologia, com notáveis impactos nos custos operacionais. Baaziz considera o caso apresentado como prova da rentabilidade da inteligência competitiva no domínio da inovação incremental, em que a abertura da ciência pode constituir um elemento revolucionário com múltiplos impactos sociais e económicos. A rentabilidade em causa manifestou-se não só nos impactos económicos a jusante, conforme descritos neste caso, mas também a montante, na rentabilização do financiamento público e privado à Investigação e ao Desenvolvimento Tecnológico.

Ferramentas para mineração e mapeamento de dados de patentes

Esta abordagem de abertura coloca exigências ao modo como a sociedade perspetiva atualmente o financiamento científico e o próprio ciclo da inovação. O potencial valor acrescentado do mapeamento de grandes quantidades de dados (a chamada *big data*) encerrados no conhecimento já alcançado e já codificado para fins específicos é, em si mesmo, um fator não exatamente quantificado no contexto dos sistemas nacionais de inovação. A mineração de dados constitui, pois, um campo científico de importância crescente no contexto interdisciplinar das ciências computacionais, no que a mineração de patentes, especificamente, oferece um ramo suplementar à geração e valorização do conhecimento, com potencialidades assinaláveis quer a nível científico, quer a nível social.

Sérgio Matos, da Universidade de Aveiro, apresentou os recentes desenvolvimentos daquela instituição a nível dos aplicativos informáticos já disponibilizados para mineração de patentes na área biomédica, nomeadamente com vista à descoberta de novo conhecimento sobre fármacos que pode emergir do conhecimento patenteado.

No caso particular da área da saúde, a documentação que pode ser “cruzada” com vista a variadas aplicações e, no limite, submetida a mineração para fins de identificação de padrões é imensa, desde literatura científica a informação clínica, a informação de patentes na área farmacêutica, a documentação explicativa sobre drogas e medicamentos, etc. Uma análise do sítio da *internet* do *European Patent Office* (EPO Espacenet) permite verificar que, até 2015, foram concedidas mais de 90 milhões de patentes na área da química. A mineração de patentes neste caso permite encontrar vias e mecanismos terapêuticos, novas classes farmacológicas e terapêuticas, novas utilizações clínicas de químicos já existentes, novas formulações ou estruturas químicas ainda não identificadas, entre várias outras possibilidades. Resultados a este nível podem, em consequência, fazer-se sentir a vários níveis, desviando potencialmente segmentos de mercado das patentes ou coberturas territoriais de foco, por exemplo, não sendo, pois, negligenciável o seu interesse para a Indústria.

É, assim, de crucial importância que os sistemas de tratamento da informação disponíveis apresentem um elevado nível de precisão e sensibilidade por forma a garantir uma forte capacidade de associação/relação linguística em aspetos como similaridades textuais, associações explícitas e implícitas, similaridades do ponto de vista semântico, etc. Para tal, é crítico que sejam apetrechados de um crescente enriquecimento semântico, para além de um algoritmo estatístico que os configure como fiáveis decodificadores de detalhes técnicos de codificação textual, tais como abreviaturas e outros aspetos característicos das diversas áreas temáticas.

A aplicação informática “Neji”, de utilização aberta, desenvolvido pela Universidade de Aveiro e orientado para a área da bioinformática, oferece diversos módulos personalizáveis para

processamento de texto e reconhecimento de conceitos, sendo enriquecido a nível semântico pelos próprios utilizadores. Apresenta resultados com diferentes níveis de fiabilidade, dependendo das vertentes temáticas da literatura analisada.

As funcionalidades das aplicações em causa passam necessariamente por uma categorização e “etiquetagem” dos termos técnicos ao longo de cada texto analisado, conforme assegurado pelo programa “Becas”, da Universidade de Aveiro, desenvolvido para identificação de conceitos. A plataforma colaborativa “Egas”, aberta e baseada na *web*, orientada para textos da área biomédica, foi também desenvolvida por um consórcio formado pela Universidade de Aveiro e pela empresa BMD, para permitir a curadoria colaborativa desses termos, criando anotações manuais e automáticas de conceitos e de relações entre os mesmos, por meio de uma interface de utilização simples. As funcionalidades de visualizações gráficas sobre análises semânticas, quer para ilustrar a proximidade de conceitos, quer para estruturação semântica de termos com base na incidência dos mesmos sobre textos associados a categorias específicas, têm provado ser também de grande utilidade.

O consórcio acima mencionado criou ainda o sistema aberto “COEUS” como estrutura simplificada de *back-end* para rápido desenvolvimento de aplicações *web* a nível de semântica. Este sistema é suficientemente flexível para integrar contributos e recursos de utilizadores com diferentes aplicativos de origem e recolher dados distribuídos de elevada heterogeneidade. O COEUS permitiu já criar o portal científico Diseasecard para exploração de semântica associada às doenças raras.

A ferramenta Patent-to-Net (P2N) foi desenvolvida por David Reymond com vista à exploração e visualização de dados de patentes registados na plataforma Espacenet. Trata-se de um aplicativo de acesso aberto desenvolvido sob licença livre, que permite a utilização de diferentes instrumentos disponíveis. O processo de tratamento de dados P2N envolve três passos sequenciais – recolha de dados e informação de patentes; exploração; e apresentação de resultados. A fase de exploração passa pela construção de matrizes dinâmicas cruzadas, matrizes de correlação e diagramas descritivos, cartografia de dados, classificação, quantificações associadas ao texto e mapeamento de conceitos. A ferramenta P2N é flexível no formato de exportação dos resultados de análise.

Tal como referido anteriormente, a análise textual permite a classificação e a associação de termos e conceitos associados às invenções patenteadas. A frequência terminológica de ocorrência nos textos é o critério mais direto, porém outros critérios de proximidade e classificação temática são contemplados.

A análise e o tratamento dos metadados associados às patentes – título, inventor, agente legal, processo de submissão, data de submissão e de publicação da patente, classificação, citações, referências, acesso e interesse gerado, etc. – são igualmente de valiosa utilidade para tratamento estatístico, nomeadamente para mapeamento de países de origem, mercados alvo, mercados financeiros de interesse (através dos pedidos de acesso

gerados), entre outros aspetos relevantes.

A ferramenta P2N, cujo desenvolvimento foi iniciado apenas em 2013, apresenta ainda algumas limitações, de carácter informático, e tem funcionalidades acrescidas já planeadas, mas ainda por desenvolver. Por outro lado, não é vocacionada para grandes quantidades de dados na área da *internet* das Coisas ou de sinais (e.g. dados CERN) ou sensores de qualquer tipo. Apresenta, no entanto, vantagens claras ao nível da integração de várias competências de tratamento de informação (incluindo de visualização) e da sua flexibilidade em termos de formatos de exportação de resultados, passíveis de serem integrados noutras cadeias de processamento e utilizados noutras vertentes de pesquisa, pelo que a sua usabilidade excede largamente o âmbito da informação de patentes. Tem grande potencial de aplicação para recolha e análise rápida de informação para estados da arte sobre determinada tecnologia, e para vários propósitos de análise de dados em desenvolvimento científico, bem como para objetivos de inovação frugal assistida por patentes.

Conclusões

Numa altura em que o conceito de inovação aberta ganha crescente interesse por parte da comunidade científica e industrial nas sociedades europeias, casos de sucesso que mostrem, por um lado, as suas potencialidades e que, por outro, ajudem a identificar os desafios colocados ao nível de modelos de financiamento do ciclo da inovação, de proteção

de dados e de delimitação de impactos são valiosos para a discussão pública. A inovação incremental com base em informação de patentes constitui um desses casos, de carácter frugal e a baixo custo, com aplicações relevantes em desafios sociais como os da área da saúde e da medicina ou outros com relevância para a indústria.

Os casos de sucesso apresentados no painel aqui descrito são, por isso, um importante testemunho dos benefícios que justificam a alteração sistémica do ciclo de inovação conforme hoje em dia o representamos, capaz de acomodar as ramificações que esta nova forma de inovação introduz, explorando as amplas possibilidades de multiplicação de impactos que o conhecimento já anteriormente gerado, e codificado para fins específicos, encerra. A informação de patentes oferece, pois, múltiplas oportunidades de transferência de tecnologia que importa aproveitar devidamente.

Recentes desenvolvimentos na área das tecnologias computacionais oferecem uma resposta eficaz, e continuamente otimizada, para esse fim, nomeadamente ao nível do tratamento de grandes quantidades de dados e informações registadas nos repositórios de patentes. Estas têm-se provado ferramentas úteis ao serviço da crescente importância da inovação frugal.

Assim, e conforme foi manifesto no painel em causa, o conhecimento científico, técnico e industrial registado nas patentes constitui uma relevante componente em emergência nos sistemas de inovação que as estratégias nacionais e comunitárias de inovação aberta em desenvolvimento não devem, pois, deixar de contemplar.

Bibliografia

1. Baaziz, A. (2016), *Transferência de tecnologia com informação aberta: Optimização de desempenho uma broca pela combinação de tecnologia de engenharia reversa e informação de patentes*. In: *Big data*, desenvolvimento sustentável e a ciência aberta. Instituto Francês, Instituto de Higiene e Medicina Tropical; Lisboa, Portugal (Junho 2016)
2. Barthez, M. D. (2016), *Caso de inovação frugal com informação aberta*. In: *Big data*, desenvolvimento sustentável e a ciência aberta. Instituto Francês, Instituto de Higiene e Medicina Tropical; Lisboa, Portugal (Junho 2016)
3. Matos, S. (2016), *Mineração de patentes para descoberta de conhecimento sobre fármacos*. Mineração de patentes para descoberta de conhecimento sobre fármacos
4. Reymond, D. (2016), *Visualisations du big data*. In: *Big data*, desenvolvimento sustentável e a ciência aberta. Instituto Francês, Instituto de Higiene e Medicina Tropical; Lisboa, Portugal (Junho 2016)
5. Moreira, B.; Saad, D.; Feldhaus, D.; Pereira, G.; Mattioli, M. (2008), *As oportunidades e os desafios do Open Innovation no Brasil*. Instituto de Inovação – Centro de conhecimento; Brasil, 2008.
6. NESTA (2016). *Better for less: Does Europe need frugal innovation?* - http://www.nesta.org.uk/blog/better-less-does-europe-need-frugal-innovation?utm_source=Nesta+Weekly+Newsletter&utm_campaign=211cda1acf-Nesta_newsletter_17_9_168_12_2016&utm_medium=email&utm_term=0_d17364114d-211cda1acf-181240777#sthash.j3NQIF1X.dpuf