

CIÊNCIA CIDADÃ



O envolvimento
popular em
atividades científicas

O crescente volume de dados gerados por instrumentos científicos de medição – como, por exemplo, a imensa quantidade de imagens do espaço exterior produzidas automaticamente por telescópios – desafia técnicas tradicionais de processamento. Alguns cientistas que analisam esses dados vêm recorrendo à ajuda de voluntários sem treinamento que, coletivamente, conseguem processá-los e também descobrir fenômenos previamente desconhecidos ali presentes. Esse tipo de colaboração é conhecido como *ciência cidadã* e tem sido usado com sucesso em diferentes áreas de pesquisa científica.

Marinalva Dias Soares

*Programa de Pós-graduação em Computação Aplicada,
Instituto Nacional de Pesquisas Espaciais*

Rafael D. C. Santos

*Laboratório Associado de Computação e Matemática Aplicada,
Instituto Nacional de Pesquisas Espaciais*

V

ários projetos científicos atuais envolvem esforços de coleta e análise de dados em grande escala, que atingem às vezes volume medido em *gigabytes* (volume de dados que cabe em um disco digital – DVD) ou *terabytes* (mais de 200 DVDs cheios). Um projeto desse tipo é a Pesquisa Sloan Digital do Céu (SDSS, na sigla em inglês), iniciado em 2000 com o objetivo de mapear objetos astronômicos como estrelas e galáxias em grande detalhe. Esse projeto usa um telescópio dedicado apenas a ele, além de computadores e programas (*softwares*) de alto desempenho para processar as imagens, incluí-las em bancos de dados e disponibilizar tais dados em diversas modalidades, tanto para astrônomos e astrofísicos quanto para o público em geral. Entre os dados coletados pelo SDSS está um catálogo com informações sobre mais de 930 mil galáxias (incluindo imagens das mesmas).

O que pode ser feito com todos esses dados? Em forma bruta, essa enorme quantidade de dados tem valor científico, mas este só será revelado quando eles forem analisados, permitindo obter informações relacionadas ao estudo que motivou a coleta. No caso do SDSS, algumas informações sobre as galáxias podem ser determinadas ou calculadas automaticamente, por programas computacionais mais simples, diretamente a partir das imagens, mas outras exigiriam maior esforço computacional ou humano para sua obtenção.



Um exemplo está no grande interesse dos astrônomos pela forma das galáxias, que podem dar indícios de como se formaram, e por outras características que não podem ser extraídas automaticamente de forma confiável – resultados obtidos por observadores humanos são considerados muito superiores. Chris Lintott, astrônomo da Universidade de Oxford (Inglaterra), estimou que um estudante de pós-graduação precisaria de três a cinco anos de trabalho incessante, sem dormir e dedicando-se integralmente a essa tarefa, para indicar manualmente as formas das galáxias contidas na base de dados do projeto. Já Alex Szalay, da Universidade Johns Hopkins (Estados Unidos), observou que provavelmente não existem astrônomos em número suficiente, no mundo inteiro, para analisar todos os dados do SDSS.

Como obter, então, as formas de quase um milhão de galáxias? Os dados brutos estão disponíveis, mas métodos automáticos não são confiáveis o suficiente para obter as informações desejadas e não é possível exigir de especialistas que dediquem o tempo necessário para o processamento manual desses dados.

Uma solução aparentemente ingênua foi usar voluntários não treinados para ajudar na classificação das formas dessas galáxias. Para isso, foi criado um projeto chamado Galaxy Zoo, que usa os dados do SDSS e os apresenta em uma página na internet, onde usuários podem opinar sobre os formatos das galáxias presentes nas imagens. Não existem restrições aos possíveis colaboradores do projeto: os usuários interessados não precisam ter conhecimentos em astronomia, mas devem se cadastrar na página e ler um breve e simples conjunto de instruções antes de começar a participar.

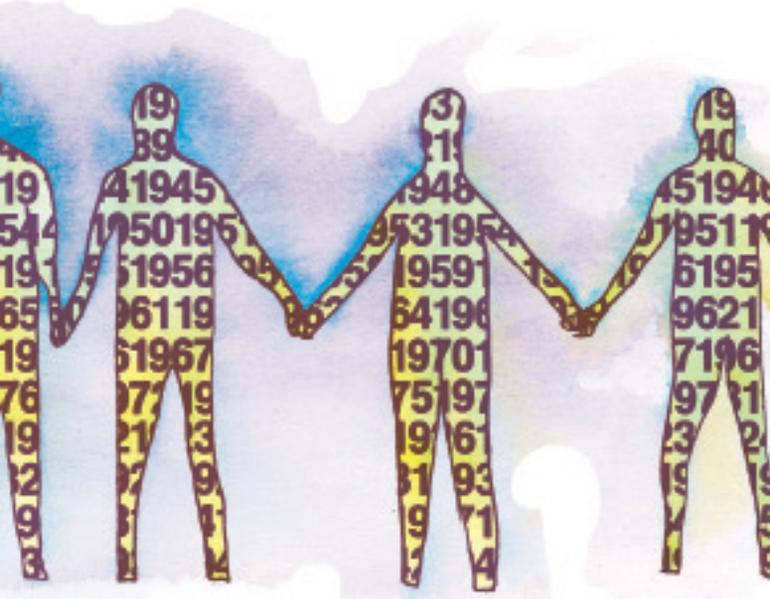
A ideia parece, à primeira vista, destinada ao fracasso. Como esperar que voluntários não treinados para o trabalho científico executem

uma tarefa específica e não trivial, após ler um simples manual de instruções? Como confiar nos resultados obtidos com a colaboração dessas pessoas? Por que alguém seria voluntário para colaborar em uma tarefa desse tipo?

Os resultados do Galaxy Zoo foram surpreendentes. A página foi aberta ao público em julho de 2007 e, em três semanas, 80 mil usuários se cadastraram e realizaram mais de 10 milhões de tarefas de classificação de imagens de galáxias. Ao final de um ano havia 150 mil usuários, que classificaram mais de 50 milhões de imagens: em média, cada galáxia foi classificada mais de 50 vezes. Ferramentas auxiliares, como páginas interativas de comentários (*blogs*) e um fórum de discussões, foram criados para que os voluntários pudessem fazer observações, sugerir objetos interessantes e rever suas colaborações.

Os dados gerados pelas colaborações desses usuários também surpreenderam. Muitas descobertas surgiram a partir de classificações feitas por voluntários. Entre esses achados estão os de que um terço das galáxias vermelhas





tem formato espiral (a maioria das espirais conhecidas é azul, o que indica a existência, nessas galáxias, de muitas estrelas recém-formadas), de que há mais galáxias azuis elípticas do que previamente estimado, de que galáxias espirais não têm uma preferência por girar no sentido horário ou anti-horário e outros. Muitos artigos científicos foram publicados com base nos resultados das classificações obtidas nesse projeto.

Um dos mais interessantes resultados surgiu a partir da observação atenta de uma colaboradora amadora: em 2007, uma professora holandesa, Hanny van Arkel, percebeu uma mancha em uma imagem do projeto e chamou a atenção de outros participantes para esse fato. De início, ninguém soube identificar o objeto, e várias teorias foram propostas, o que levou a observações mais detalhadas. A

mancha foi identificada, em 2010, como uma nuvem de gás ionizada pela radiação de um buraco negro de uma galáxia vizinha – um objeto raro, descoberto por uma participante amadora do Galaxy Zoo. Uma das conclusões dos criadores desse projeto é que computadores podem ficar cada vez melhores na classificação de galáxias, mas a capacidade de olhar uma imagem e perguntar “que objeto estranho é este?” é inerentemente humana.

Ciência de não cientistas _O projeto Galaxy Zoo é um dos exemplos mais célebres de ciência cidadã, mais conhecida pelo nome em inglês, *citizen science*. Esse termo é empregado para descrever, de forma geral, projetos com finalidades científicas que usam não cientistas como voluntários. Projetos de ciência cidadã são diferentes de estudos aparentemente similares, como o Seti@home (análise de sinais de rádio em busca de indícios de vida extraterrestre inteligente) ou o Einstein@home (busca por objetos astronômicos com intensa força gravitacional, como pulsares, buracos negros, estrelas de nêutrons e outros), que envolvem análises automáticas e apenas utilizam o tempo ocioso de computadores pessoais. Na ciência cidadã, os voluntários participam ativamente dos projetos, cedendo tempo e trabalho.

A ciência cidadã também é considerada diferente da distribuição de tarefas por voluntários, pagos ou não por seus serviços (processo conhecido, em inglês, como *crowdsourcing*). Os projetos de ciência cidadã têm base tradicionalmente científica e não comercial. Os cientistas envolvidos nesses projetos consideram os voluntários colaboradores e aprendizes e os voluntários podem participar na coleta de dados de interesse científico ou na análise desses dados.



The Cornell Lab of Ornithology
Exploring and Conserving Nature

About Us | What We Do | Give | All About Birds

MISSION: CITIZEN SCIENCE

What We Do
 Research
 Education
 Citizen Science
 Technology
 Conservation
 Lab Programs

How YOU CAN HELP
 We rely on your support to conduct to further our mission to understand birds and other wildlife, to involve the public in scientific discovery, and to use our knowledge to protect our planet.

DONATE NOW

Sign up for eNews

Each day, bird watchers report tens of thousands of bird observations to citizen-science projects at the Cornell Lab of Ornithology, contributing to the world's most dynamic and powerful source of information on birds. The Cornell Lab has been at the forefront of citizen science since 1962, and continues to innovate by creating online tools enabling people to share and explore their data. Today, the online community taps into millions of records to see how their own sightings fit into the continental picture, and scientists analyze the data to reveal striking changes in the movements, distribution, and numbers of birds throughout time.

More than 200,000 people contribute to the Cornell Lab's

CITIZEN SCIENCE

Our projects suit people of all ages and experience levels.

QUICK FACTS

- More than 60 scientific papers have used Cornell Lab citizen science data since 1997
- 87,000 checklists reported from the first eBird Checklist Report Card in 2010
- More than 90,000 eSightings

Centenas de milhares de pessoas, de todo o mundo, participam de projetos de ciência cidadã pela internet, como o Galaxy Zoo (na área de astronomia), o Coral Watch (na área de biologia marinha) e o Cornell Lab of Ornithology (na área de ornitologia)



Ciência cidadã não é um conceito novo – um dos projetos mais antigos, iniciado em 1900, é a contagem de pássaros, hoje coordenada pela sociedade Audubon, uma organização ambiental não governamental dos Estados Unidos. Grupos de voluntários contam pássaros em regiões pré-definidas, obedecendo a algumas regras ditas pela sociedade, com o objetivo de fazer um censo simplificado sobre os tipos e populações de pássaros. Atualmente, a aplicabilidade desse tipo de projeto vem aumentando graças à redução dos custos e à ampliação do acesso a telecomunicações – em particular, do acesso à internet e a redes de dados por meio de telefonia celular.

Com os recursos tecnológicos atuais, é simples criar um projeto a partir de um repositório de dados acessível pela internet. A mesma estrutura usada para armazenar os dados pode ser empregada para disponibilizar os que devem ser processados, fornecer informações aos participantes e coletar os resultados de suas análises. O que pode ser difícil é motivar as pessoas a agir como cidadãos-cientistas.

Entender essa motivação – em muitos casos, os voluntários reduzem seu envolvimento depois de um período inicial de experimentação – é crucial para o gerenciamento dos projetos de ciência cidadã. Por isso, cientistas envolvidos em algumas dessas iniciativas fizeram estudos sobre tais colaboradores. Um desses estudos mostrou que, em certas circunstâncias, os resultados ob-

tidos por cidadãos-cientistas podem ser melhores que os gerados por técnicos pagos e bem treinados, justamente por causa da motivação dos voluntários, que se sentem recompensados apenas com a sensação de fazer parte de algo grande e útil.

Estudos informais feitos pelos autores deste artigo com um grupo de voluntários, em um projeto de identificação de regiões urbanas em imagens de satélites (ver ‘Interação investigada’), sugere também fatores pessoais na motivação dos voluntários, que mostram maior interesse em colaborar com projetos quando sua participação é solicitada por conhecidos, em comparação com pedidos anônimos por meio da internet.

O sucesso de alguns projetos de ciência cidadã, confirmado não só pelo número de publicações decorrentes dos resultados das análises feitas pelos voluntários, mas também pelo grande envolvimento de pessoas (sabe-se que indivíduos ativos em alguns projetos tendem a colaborar em iniciativas semelhantes), levou ao agrupamento de projetos semelhantes com a intenção de compartilhar recursos, em especial a base de usuários já cadastrados, para solicitar participação em tarefas semelhantes. Dois desses grupos de projetos de ciência cidadã são o Zooniverse e o Cornell Lab of Ornithology.

O Zooniverse é um portal, na internet, que hospeda vários projetos de ciência cidadã (vários na área da astronomia e um sobre o clima da Terra). Esses projetos são

INTERAÇÃO INVESTIGADA

O UrbanZoo, projeto inspirado no Galaxy Zoo e desenvolvido pelos autores, pede que voluntários escolham uma classe (ou um rótulo) para definir regiões aparentemente uniformes presentes em uma imagem de satélite de alta resolução que mostra uma área urbana. Diferentemente de alguns projetos tradicionais de ciência cidadã, o objetivo principal não é obter uma classificação precisa das regiões apresentadas, mas coletar informações sobre a interação dos voluntários com o programa de computador que apresenta as tarefas de classificação.

De modo proposital, foi escolhido um problema complexo, com dados incompletos e conflitantes, para verificar se a opinião coletiva de usuários com diferentes perfis pode ajudar a identificar características e problemas do processo original que gerou os dados. Os estudos estão em andamento, mas resultados preliminares indicam que as opiniões coletivas dos usuários podem servir para rotular a imagem e para identificar regiões que não podem ser facilmente classificadas, para revelar conflitos e inconsistências na seleção dos rótulos usados na tarefa e até para indicar possíveis erros cometidos por especialistas na análise dessas imagens.

mantidos e desenvolvidos pela Citizen Science Alliance (CSA), uma colaboração entre cientistas, programadores de computação e educadores de vários países, que coletivamente desenvolvem, gerenciam e utilizam (com a participação de universidades e instituições de pesquisa de vários países) projetos de ciência cidadã baseados na internet. Segundo a página do Zooniverse na internet, mais de 380 mil voluntários atuam em seus projetos.

Já o Cornell Lab of Ornithology, ligado à Universidade Cornell (Estados Unidos), é líder mundial no estudo, apreciação e conservação de aves. Esse laboratório, em que cientistas, ambientalistas, engenheiros, educadores e estudantes trabalham em conjunto, tem como objetivos compreender as aves e outros animais selvagens, envolver o público na descoberta científica e usar o conhecimento adquirido para proteger o planeta. Mais de 200 mil voluntários participam de seus projetos de ciência cidadã, iniciados em 1966.

Os dados são confiáveis? Alguns cientistas não acreditam que os dados coletados ou processados por voluntários não-treinados tenham qualidade, e até criticam projetos de ciência cidadã como sendo válidos apenas para divulgação ou educação. Outros afirmam que o consenso científico não deve levar em conta opiniões de não cientistas.

Essas opiniões não têm fundamento porque, nos projetos sérios de ciência cidadã, que contam com observações e/ou opiniões de voluntários, as colaborações destes são sempre filtradas ou reavaliadas. Isso pode ser feito de modo direto (com o uso, por exemplo, de métodos que avaliam o consenso entre as opiniões diversas dos voluntários) ou indireto (procurando, por exemplo, confirmar evidências obtidas pelos não cientistas por meio de dados coletados de outra forma). Essa última modalidade é usada em projetos de coleta de dados naturais como o *Coral Watch*, que monitora as condições dos recifes de coral, ligado à Universidade de Queensland (Austrália)

Mesmo em projetos onde a principal fonte de informações está nas opiniões dos voluntários, ao final a tomada de decisão é dos cientistas. Em projetos que envolvem classificações feitas pelos voluntários, é possível, por exemplo, selecionar somente dados classificados por certo número de voluntários que apresentem algum grau de concordância – esses critérios quantitativos podem ser definidos pelo cientista que usará os dados em suas pesquisas. O projeto UrbanZoo, criado pelos autores, permite também identificar objetos que receberam muitas rotulações ou opiniões divergentes, marcando-os, na base de dados, como merecedores de atenção específica por um especialista que supervisiona o projeto. Assim, o especialista pode atuar como mediador de opiniões em casos de falta de consenso, o que exige muito menos tempo e esforço desse pesquisador do que se ele tivesse que rotular todos os objetos da base.

É importante lembrar que voluntários de projetos de ciência cidadã colaboram com os cientistas organizadores, mas de forma alguma os substituem. Os dados coletados podem fornecer subsídios para conclusões científicas, mas isso depende de uma análise crítica feita por um cientista. É inegável, porém, que alguns projetos têm tido enorme sucesso, não só quanto à motivação dos voluntários, mas também quanto ao aproveitamento, sob condições cuidadosas, dos dados coletados para fins científicos. 

Sugestões para leitura

HAND, E. 'Citizen science: people power', in *Nature*, n° 466, p. 685, 2010 (disponível em <http://www.nature.com/news/2010/100804/full/466685a.html>)
 SOARES, M. D.; SANTOS, R.; VIJAYKUMAR, N. L. & DUTRA, L. V. 'Citizen science-based labeling of imprecisely segmented images: case study and preliminary results', em *VIII Simpósio Brasileiro de Sistemas Colaborativos*, Belo Horizonte, p. 87, IEEE Computer Society, 2010 (disponível em <http://www.computer.org/portal/web/csdl/doi/10.1109/SBSC.2010.20>)

NA INTERNET

>> Citizen Science Alliance: <http://citizensciencealliance.org/>
 >> Zooniverse Projects: <http://www.zooniverse.org/>
 >> Urban Zoo: <http://www.lac.inpe.br/UrbanZoo>