

# COMPARTILHAMENTO E INTEGRAÇÃO DE DADOS

## A CONSTRUÇÃO DO CONHECIMENTO SOBRE BIODIVERSIDADE

**Debora Pignatari Drucker**

**Flávia Fonseca Pezzini**

O conhecimento sobre a biodiversidade brasileira é fundamental para o sucesso na sua conservação e uso sustentável. Acompanhar as mudanças na diversidade e na distribuição das espécies e suas relações com o ambiente permite definir estratégias e mecanismos efetivos para a conservação. Mas como são obtidas as informações que permitem conhecer cada espécie? Estimativas apontam que a biota brasileira representa mais de 20% da biota mundial. Diante de tamanha grandiosidade, como os pesquisadores fazem para conhecer a biodiversidade e seus processos? Como podem analisar se uma determinada atividade econômica está realmente sendo sustentável?

A resposta para essas e outras perguntas, por incrível que pareça, é a mesma: eles fazem isso integrando e compartilhando

dados. Após visitarem densas florestas, caminharem por campos abertos ou investigarem qualquer outro ecossistema, depois de consultarem imagens de satélites, bases de dados e outros instrumentos, os pesquisadores voltam aos seus laboratórios com os dados sobre o que viram e as amostras que recolheram, quando necessário. Essas dados são registrados, organizados e analisados de maneira específica de acordo com o que se quer estudar, resultando em informações. Os materiais são processados e analisados em laboratórios ou em coleções científicas.

Dados são a base da pesquisa científica em todos os campos do conhecimento, não apenas nas áreas relacionadas à biodiversidade. São peças fundamentais para embasar o avanço da ciência e informar a tomada de decisões. Podem ser obtidos

de diferentes maneiras, em diferentes locais e analisados e reanalisados mesmo muito tempo depois de sua obtenção.

Pesquisas em biodiversidade são fortemente baseadas em representações quantificáveis, ou seja, números que representam entidades ou fenômenos, e estimativas. Uma pesquisa pode incorporar dados coletados por pesquisadores em diferentes locais, assim como dados publicados em repositórios ou na literatura para serem reanalisados em novos estudos. Os dados podem também ser integrados de modo a construir um conhecimento complexo. É possível, por exemplo, integrar dados de organismos diferentes que ocorrem em um mesmo local, como dados sobre a estrutura da vegetação e a diversidade de aves em uma unidade de conservação. É possível combinar dados de um mesmo grupo em tempos diferentes, como a variação da composição da comunidade de aves dessa unidade de conservação em diferentes estações do ano. Ou ainda, combinar informações de diferentes grupos biológicos em diferentes locais e tempo – para descobrir, por exemplo, como a comunidade de aves varia ao longo do ano em áreas com diferentes estruturas de vegetação.

Entretanto, para que as diferentes informações sejam combinadas é necessário que os diversos grupos de pesquisa considerem pontos tais como: precisam saber da existência dos diversos conjuntos de dados. Depois, precisam ter acesso a eles. E, para utilizá-los, precisam entender como foram obtidos e se são dados brutos, ou seja, se são apresentados assim como foram coletados em campo, antes de qualquer análise. Precisam ainda identificar os responsáveis pelos dados, bem como os direitos de uso e redistribuição dos mesmos, de modo a assegurar o devido crédito aos autores e viabilizar novas parcerias científicas a partir de colaborações em novas análises e interpretações e na publicação de resultados.

Com as novas tecnologias computacionais, a quantidade de dados disponíveis para a pesquisa tem aumentado de forma acelerada. Ao mesmo tempo, o avanço nas tecnologias de comunicação tem tornado a ciência cada vez mais colaborativa. Portanto, cada um desses passos – aquisição, gestão, compartilhamento e divulgação de dados científicos –, e especialmente a boa documentação de um conjunto de dados, tem cada vez mais importância. Dados de campo têm alto valor, exigem esforço e re-



---

Oficina de gerenciamento de dados e metadados em ecologia oferecidos pelo PPBio contribuem para a divulgação da importância e do cuidado com a qualidade e do compartilhamento de dados

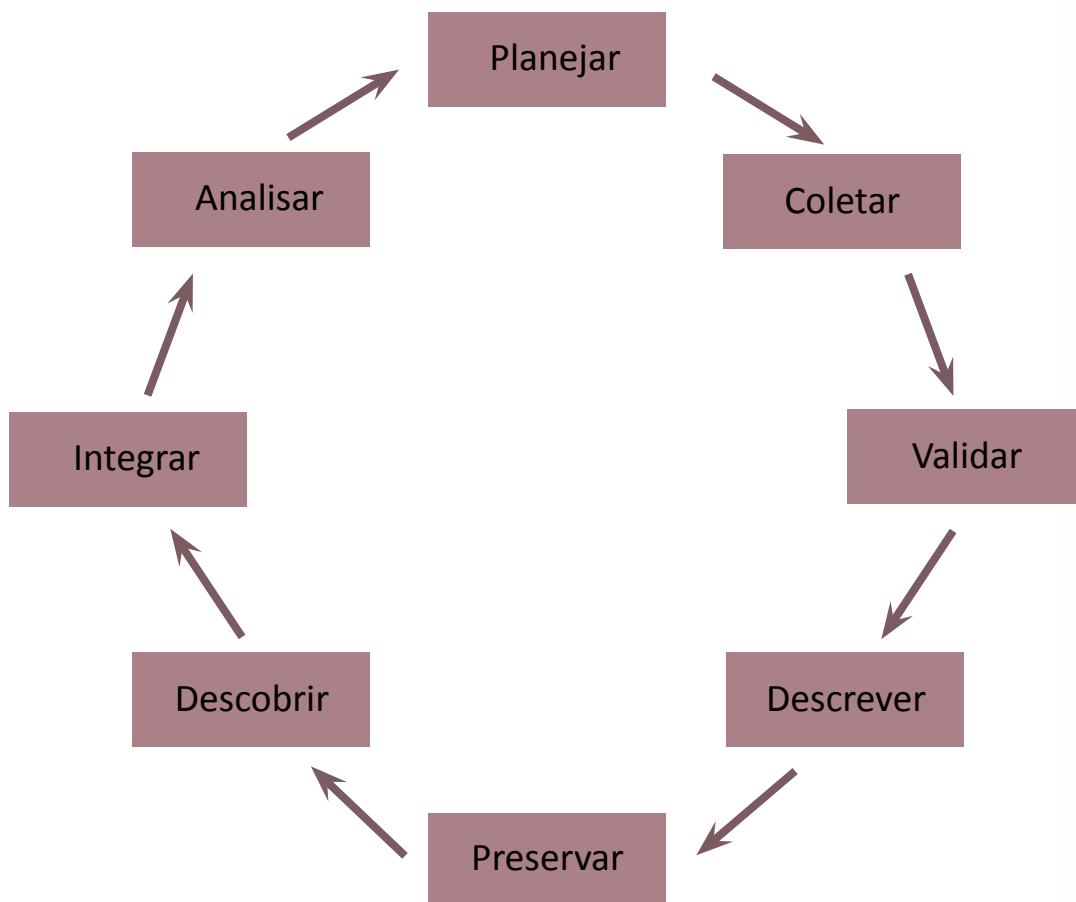
curso para sua obtenção que não podem ser desperdiçados.

### **Dados sobre biodiversidade**

Há inúmeros tipos de dados que são coletados continuamente pelos cientistas para representar a biodiversidade: dados taxonômicos, relacionados à classificação de seres vivos em grupos como espécies, gêneros e famílias; dados ambientais, como aqueles sobre o solo e o clima; dados

ecológicos, sobre as interações entre espécies, por exemplo; dados de distribuição geográfica de organismos; dados de experimentos de campo, laboratório ou casa de vegetação, entre outros. Os levantamentos realizados para obter esses dados são planejados com meses de antecedência e, na maioria das vezes, direcionados ao estudo de um grupo biológico ou um fenômeno ecossistêmico pré-determinado.

De acordo com as características das perguntas que os pesquisadores querem responder, são definidos os instrumentos de medição, a composição da equipe, os materiais para a preservação das amostras, a estrutura da tabela a ser preenchida no campo, os dispositivos eletrônicos para a gravação, dentre outros itens importantes para o sucesso de uma expedição. É também recomendável que o planejamento do trabalho de campo incorpore aspectos sobre a organização e o uso futuro dos dados. Isto significa considerar todo o ciclo de vida dos dados – um contínuo de etapas de curadoria necessárias para transformá-los em novo conhecimento. Em uma perspectiva ampla, curadoria de dados científicos é entendida como o conjunto de atividades e processos realizados para gerenciar, manter, validar e armaze-



---

Ciclo de vida dos dados, desde seu planejamento e geração até sua preservação, integração e análise, produzindo resultados que levarão ao avanço do conhecimento e estimularão novas perguntas, levando a um novo ciclo.

Fonte: adaptado de <https://www.dataone.org/best-practices>

nar dados de pesquisa de forma a preservá-los e permitir seu reuso.

O ciclo de vida dos dados abrange etapas que vão desde a geração e aferição de sua qualidade, a documentação (com metadados apropriados) e preservação,

até sua recuperação por buscas e consultas, sua integração, análise e publicação dos resultados. Como a construção do conhecimento é contínua, uma volta no ciclo produzirá resultados que incitarão a geração de novas perguntas científicas, as

quais levarão ao início de um novo ciclo de planejamento de novas investigações para respondê-las. A disponibilização pública dos dados em repositórios científicos aumenta muito o aproveitamento e o valor social das pesquisas, uma vez que possibilita o cruzamento de dados produzidos por diferentes grupos de pesquisa, e vem sendo cada vez mais exigida por revistas científicas e agências de financiamento.

As perguntas a serem respondidas por um grupo de cientistas podem não ser as mesmas perguntas de interesse de outro grupo. Entretanto, frequentemente, os dados brutos coletados por um grupo são úteis para outros grupos, especialmente se forem coletados em uma mesma área ou em uma mesma escala espacial. Um exemplo são os dados coletados em parcelas permanentes, que são áreas delimitadas em locais de interesse para coleta de dados e amostradas durante longo prazo e nos quais inúmeros projetos de pesquisa diferentes são conduzidos. Uma característica de parcelas permanentes é o investimento em infraestrutura e condições de pesquisa para viabilizar a investigação de fenômenos que demandam um longo período de observações e coleta de dados para serem compreendidos.

Dados publicados em repositórios ou na literatura podem ser reanalisados em novos estudos. Tecnologias computacionais facilitam a produção, o acesso e a integração de dados de procedência diversa como dados obtidos por sensoriamento remoto, frequentemente por meio de satélites; dados gerados por sensores de campo, como estações meteorológicas; ou câmeras que registram a presença de vertebrados. Ainda assim, a obtenção continuada de dados em campo por pessoas é imprescindível para produzir representações da biodiversidade confiáveis. Esses dados dependentes da observação e anotação humana caracterizam-se por seu alto valor agregado, tanto do ponto de vista do esforço empreendido para sua obtenção, bem como sob o aspecto dos recursos necessários para viabilizar sua aquisição. São heterogêneos, com variações no formato, modelo e semântica, características desafiadoras em termos de viabilizar sua preservação, recuperação e reutilização.

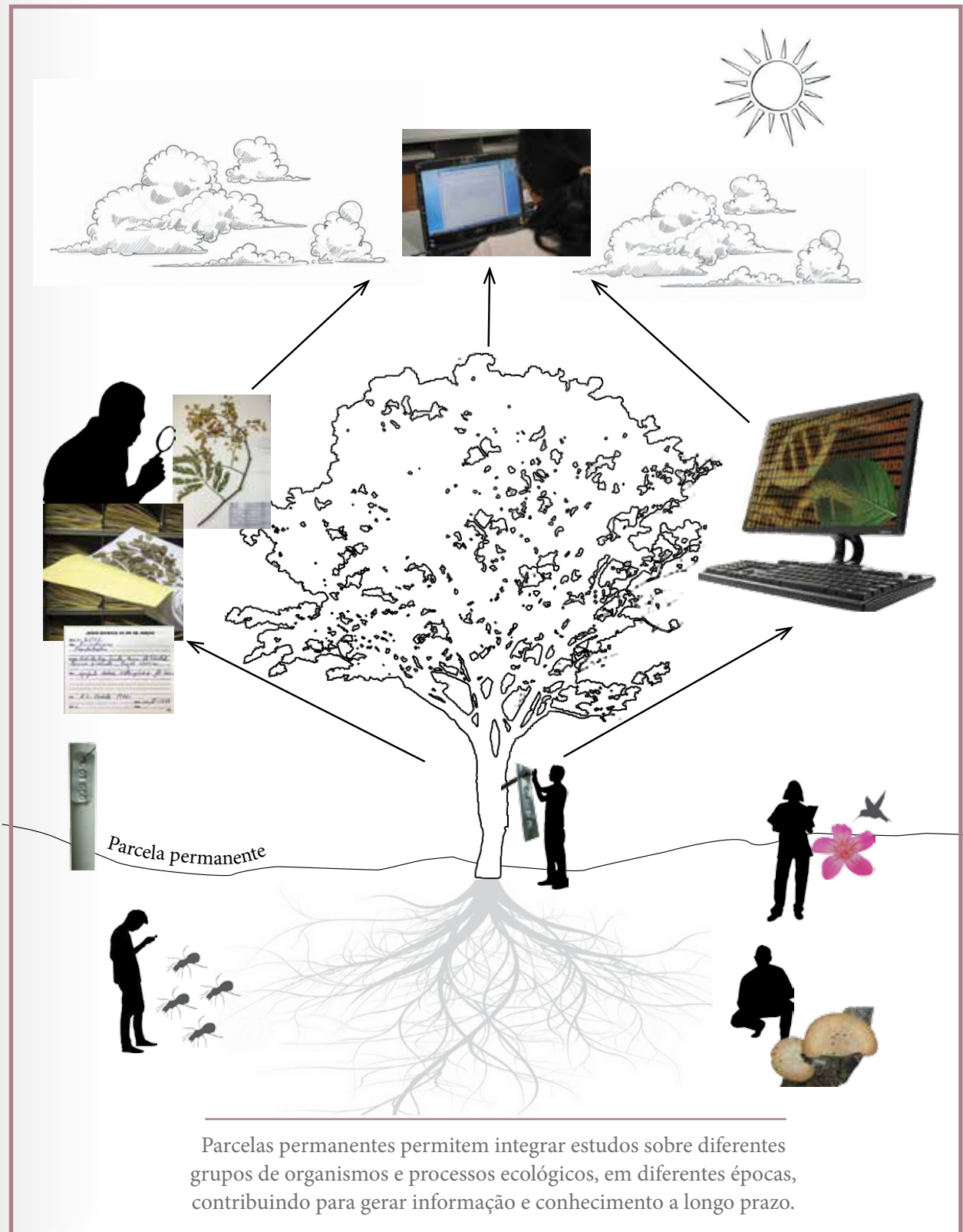
Por exemplo, quando o interesse está voltado para a compreensão da quantidade estocada e da dinâmica de carbono em uma floresta, pesquisadores utilizam uma série de técnicas em conjunto. Boa parte do carbono estocado em uma flo-

resta está armazenado nas árvores. Para estimar a quantidade estocada no componente arbóreo, é preciso medir o diâmetro e a altura das árvores de determinado local. Em geral, a parcela no campo, seja ela permanente ou temporária, é identificada e marcada com estacas e fitas. Muitas vezes, os dados de altura são estimados, por diferentes técnicas, pois é difícil visualizar o topo das árvores. No caso do diâmetro, há um protocolo a ser seguido, e convencionou-se medi-lo à altura do peito, considerado como 1,30 m do solo. As medições são anotadas em planilhas de campo, nas quais são anotadas também singularidades das árvores medidas, como o fato de algumas delas bifurcarem antes da altura de 1,30 m. Todo e qualquer detalhe é importante para o resultado final e por isso o cuidado em anotá-lo devidamente é essencial.

Como, normalmente, há o interesse em acompanhar o crescimento das árvores ao longo do tempo, um número de campo é atribuído a cada uma delas. Esses números, registrados também em pequenas placas fixadas nas árvores, passam a ser seus códigos de identidade e são informatizados em bases de dados no computador, juntamente com os dados das me-

dições de altura e diâmetro realizadas. A localização das árvores também é registrada, normalmente com os valores x e y de um plano cartesiano. Esse procedimento viabiliza que novas informações atribuídas a cada árvore sejam registradas e comparadas. Além de todo o esforço para medir as árvores acima do solo, um grande desafio no estudo da estocagem de carbono é a medição de sua presença nas raízes, um processo custoso. Para aumentar ainda mais a complexidade, há também o componente da respiração do solo para ser contabilizado nas equações de carbono. Além disso, diferentes espécies têm diferentes densidades de madeira, o que também afeta as estimativas. Para se conhecer as diferentes espécies, é preciso identificá-las, o que requer a amostragem de galhos, folhas, flores e frutos. O material coletado, chamado de testemunho ou *voucher*, é acondicionado em prensas que serão secas em estufas.

Não importa qual a pergunta ou o objetivo da pesquisa, é sempre preciso coletar o material testemunho. Trata-se de exemplares dos organismos estudados que são coletados exclusivamente para armazenamento em coleções científicas. O material testemunho pode ser uma planta (geral-



Parcelas permanentes permitem integrar estudos sobre diferentes grupos de organismos e processos ecológicos, em diferentes épocas, contribuindo para gerar informação e conhecimento a longo prazo.

mente partes contendo folhas, flores ou frutos), que será incorporada a herbários; exemplares de animais que serão incorporados à coleções zoológicas; ou mesmo tecidos, tanto de animais como de plantas, cujas coleções vêm crescendo cada vez mais em importância.

Ao coletar um *voucher*, os pesquisadores também anotam as informações associadas, chamadas tecnicamente de metadados. Essas informações incluem o local, as coordenadas geográficas e a data de coleta, o nome do coletor e seu número de coleta, informações taxonômicas, e quem identificou o indivíduo coletado. Esse último, muitas vezes, incorporado aos metadados posteriormente.

Assim que o *voucher* é depositado em uma coleção, ele recebe um número de tombo, número exclusivo que o identifica como um item daquela coleção. Em pesquisas botânicas, normalmente várias duplicatas de um mesmo indivíduo são coletadas e distribuídas por diferentes herbários, prática que pode ser mais difícil em pesquisas zoológicas. Cada duplicata recebe um número de tombo

diferente de acordo com o herbário no qual foi depositada, mas todas estarão interligadas pelo nome e número do coletor. Assim, todos os espécimes que foram estudados em um levantamento de campo possuem informações associadas nos respectivos metadados e também material testemunho depositado em coleções científicas.

O controle do nome e número de coletor é muito importante e deve ser feito pelo pesquisador. O número é um identificador único que pode ser citado como referência em artigos científicos, seja de revisão taxonômica ou estudos ecológicos. Nos últimos anos, sequências genéticas depositadas em bases de dados online, como o GenBank<sup>1</sup>, passaram a utilizar esse código como referência dos dados ali depositados. O número de coletor é também de extrema importância para rastrear e atualizar informações sobre as espécies nas diferentes coleções, bem como para associa-las aos dados coletados em campo.

Dados sobre biodiversidade são dinâmicos: com o passar do tempo é possí-

---

<sup>1</sup> O GenBank é uma base de dados de sequências genéticas amplamente utilizado, que pode ser consultada em <http://www.ncbi.nlm.nih.gov/genbank/>



vel agregar informações a eles ou corrigir erros, aumentando sua qualidade e utilidade. Aquela mesma árvore marcada em uma parcela permanente em um estudo sobre estoque de carbono, por exemplo, será estudada com outros objetivos. É possível ter, além do nome que identifica a espécie, uma série de informações, tais como: medidas de diâmetro e altura (estrutura); dados sobre crescimento (dinâmica), produção de flores, frutos e renovação de folhas (fenologia); quantificações do tipo de solo em que ela ocorre; quais animais a polinizam e se alimentam dela e quais são seus predadores (interações ecológicas); ou quais compostos químicos ela produz e para que eles podem ser úteis (bioprospecção); como ela lida com as variações de disponibilidade de luz e água (fisiologia); em quais outros lugares essa mesma espécie está presente (distribuição geográfica); há quanto tempo essa espécie se diversificou e está ali naquele ambiente (biogeografia); entre outros dados. Em geral, essas informações são obtidas ao longo de meses ou anos por meio de estudos realizados por diferentes grupos de pesquisa — muitas vezes são produto de várias dissertações de mestrado e teses de doutorado.

Um dos desafios relacionados à preservação de dados sobre biodiversidade para reuso futuro é acompanhar as mudanças que ocorrem nos sistemas de referência taxonômicos ao longo do tempo. Caso o taxonomista faça uma nova determinação, é desejável poder rastrear todos os estudos aos quais ela está ligada e atualizar seu nome. Isso inclui outras bases de dados online, como o de sequências genéticas, GenBank. Mas todas podem estar ligadas ao voucher depositado em uma coleção, o que facilita a realização de mudanças conforme elas ocorram.

### **A importância do cuidado com os dados**

O trabalho com dados requer cuidado e atenção. Rasuras e falta de informações no preenchimento de fichas de campo ocasionam problemas de interpretação nas análises futuras. Erros de digitação cometidos, por exemplo, ao criar uma tabela a partir das anotações nas fichas de campo, também devem ser evitados, e todo dado digitado deve ser conferido. Quando os dados são adquiridos por meio de sensores e armazenados em dispositivos de gravação automáticos, é preciso realizar procedimentos de verificação que auxiliem a detectar valores improváveis que



Pesquisadores trabalhando em parcelas permanentes: investimento em infraestrutura para pesquisas de médio e longo prazo integrando dados coletados por diferentes equipes

possam estar sendo registrados devido, por exemplo, a problemas de calibração de instrumentos.

Ao realizar seu experimento, um pesquisador tem claros os detalhes necessários para usar o conjunto de dados obtidos. Sabe como uma coleta foi feita, como o experimento foi montado ou o tipo, marca e modelo dos aparelhos utilizados — após realizar o experimento essas informações estão vivas em sua memória. Mas elas logo se perderão se não forem organizadas, seja no papel ou no computador. Além de ficarem indisponíveis para outras pessoas caso isso não seja feito, é muito provável que o próprio pesquisador esqueça a maioria dos detalhes em pouco tempo. A falta de registro organizado dos dados, obtidos às custas de muito esforço e com recursos financeiros e humanos, inviabiliza que eles sejam aproveitados em seu máximo potencial.

### **Computação, uma aliada**

A organização de dados sobre a biodiversidade em uma só base de dados pode ser muito complexa, dada a variedade de tipos de informações que podem ser coletadas e a dificuldade de prever quais serão coletadas no futuro. O Centro Na-

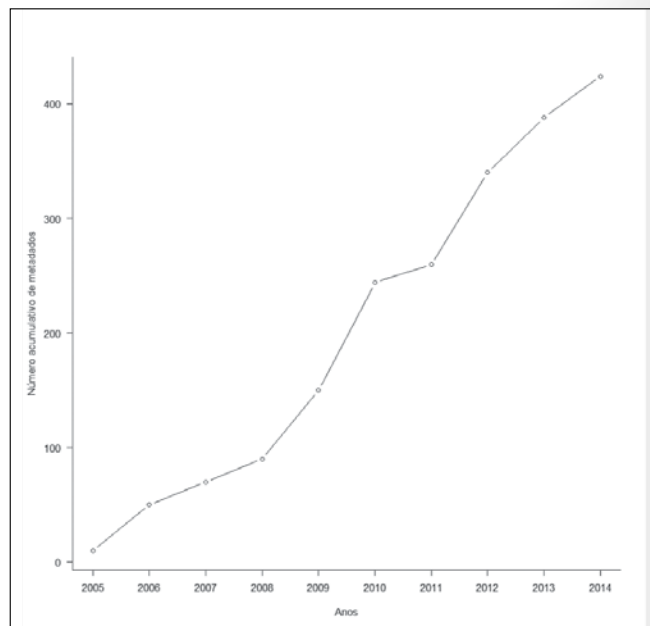
cional de Análise e Síntese Ecológica (NCEAS, na sigla em inglês), da Universidade da Califórnia, em Santa Bárbara, nos Estados Unidos, reúne pesquisadores que buscam integrar e analisar dados ecológicos para produzir estudos de síntese do conhecimento. A experiência desses cientistas mostrou que encontrar, acessar e compreender dados coletados por outros pesquisadores para poder integrá-los eram algumas das etapas mais demoradas e trabalhosas da pesquisa.

Essa constatação levou a equipe de ecoinformática do NCEAS a desenvolver ferramentas computacionais de código aberto para armazenar, organizar e preservar dados ecológicos. Uma dessas ferramentas é o Metacat, um repositório no qual dados e seus respectivos metadados são armazenados e documentados em uma estrutura e linguagem fixa, seguindo um padrão chamado “*Ecological Metadata Language - EML*”.

O Metacat é um repositório flexível e, justamente por isso, é recomendável a presença de um gestor de dados, uma pessoa que assegura que a documentação seja feita de forma a viabilizar o reuso dos dados. Em 2010, o Programa de Pesquisa em Biodiversidade Amazônia Ocidental

(PPBioAmOc) começou a usar o Metacat para armazenar dados e metadados. Desde 2005, há, no PPBioAmOc, uma pessoa exclusivamente responsável por garantir que as informações sejam providenciadas pelos pesquisadores e, posteriormente, disponibilizadas. A presença do gestor de dados foi um fator determinante para os resultados positivos do programa, colaborando para o incremento contínuo na quantidade de dados depositados no repositório. Pesquisadores, muitas vezes, precisam de orientação de como organizar e disponibilizar seus dados e raras vezes de incentivo e explicação de porque é importante. O gestor de dados auxilia o pesquisador a realizar as boas práticas ao longo do ciclo de vida dos dados, em especial nas etapas de validação, descrição e preservação, assegurando seu reuso e a reprodutibilidade de resultados. Uma política de dados bem definida é outro fator chave que vem consolidando a cultura do compartilhamento de dados no PPBio AmOc<sup>2</sup>.

Pesquisadores e gestores de dados utilizam para seu trabalho outro programa criado pelo NCEAS, uma ferramenta de edição de metadados, também no EML, chamada



Número acumulado de metadados disponíveis no repositório PPBio Am-Ac de 2005 a 2014. O incremento anual de dados descritos e organizados no sistema demonstra a adesão da comunidade de pesquisa aos procedimentos de gestão de dados.

Morpho, a qual interage com o Metacat. Essa ferramenta funciona como um guia para a estruturação de metadados em consonância com padrões abertos amplamente adotados, criando arquivos em formato XML – uma linguagem de marcação que facilita a recuperação da informação por mecanismos de busca, utilizando por exemplo palavras-chave.

Mas ainda há desafios a serem vencidos. É preciso, por exemplo, estimular a ado-

2 A política de dados do PPBio pode ser consultada em [https://ppbio.inpa.gov.br/sites/default/files/politica\\_dou.pdf](https://ppbio.inpa.gov.br/sites/default/files/politica_dou.pdf)

ção de termos padronizados, de modo a favorecer o uso de palavras-chave iguais para representar o mesmo fenômeno ou entidade. Por exemplo, o uso das palavras “sapo” e “anura” como palavras-chave em diferentes metadados pode influenciar os resultados exibidos em determinada busca. Diversos grupos vêm trabalhando, nacional e internacionalmente, com o desenvolvimento de técnicas para facilitar essas representações, favorecendo o que chamamos de interoperabilidade semântica. É também fundamental descrever os procedimentos usados para integrar os dados, os códigos e as equações utilizadas nas análises, e os métodos de visualização de dados escolhidos. Essas práticas ainda são incomuns na comunidade científica, mas provavelmente passarão a ser mais frequentes ao longo dos próximos anos, à medida que as ferramentas e tutoriais se tornem mais acessíveis.

### **O futuro da gestão e compartilhamento de dados sobre biodiversidade**

Em muitos países, os dados brutos coletados pelos pesquisadores começaram a ser reconhecidos e tratados como um dos principais produtos das pesquisas. Agências de fomento, como a “National

Science Foundation” nos Estados Unidos e o “Natural Environment Research Council” no Reino Unido, já exigem que os pesquisadores incluam um plano de gestão de dados na proposta de pesquisa a ser submetida para financiamento. Esse plano deve incluir informações como quem será o responsável pelos dados, quais os tipos de dados produzidos, padrões de dados e metadados que serão adotados, onde e como serão preservados ao longo dos anos, mesmo depois que o projeto tiver sido encerrado, quem terá acesso a esses dados, como eles serão disseminados, reutilizados e compartilhados. Além disso, importantes revistas científicas internacionais na área de biologia, como “American Naturalist” e “Heredity”, passaram a exigir que os dados relacionados aos artigos nelas publicados sejam disponibilizados em repositórios públicos.

No Brasil, algumas agências vêm sinalizando a adoção de procedimentos semelhantes aos das agências de fomento americana e britânica, o que indica que a prática de planejamento da gestão de dados passará a ser uma realidade também no contexto nacional. Essa iniciativa beneficia as pesquisas e é exemplo de transparência e bom uso de recursos públicos.

Cada vez mais há interesse em compreender de que maneira os resultados de pesquisas levam a transformações na sociedade. Métricas tradicionais de impacto de pesquisas são baseadas em citações de artigos científicos em novos artigos científicos, ou seja, avaliam o impacto na esfera da ciência. Como complemento, métodos inovadores para a medição dos impactos da ciência vêm sendo propostos por grupos que pesquisam como detectar os acessos aos resultados de pesquisas no mundo digital, uma nova área de pesquisa que tem sido chamada de altmetria, nome derivado de métricas (ou medições) alternativas.

Na altmetria, as informações sobre o acesso às informações científicas são registradas a partir da interação das pessoas (pesquisadores ou não) com portais de ciência, por exemplo, pelo registro de visualizações e downloads de conteúdos científicos, discussões em blogs de ciência, citações na Wikipedia ou em documentos voltados para políticas públicas, divulgação em mídias jornalísticas e compartilhamento, etiquetagem e comentários sobre conteúdos científicos em mídias sociais como Twitter, Facebook ou G+. Essas mé-

tricas possibilitam analisar a atenção que determinado produto de pesquisa alcançou, ou seja, quantas pessoas foram expostas ao conteúdo, sua disseminação – onde e como um resultado de pesquisa está sendo discutido ou compartilhado, o impacto e sua influência, tanto entre cientistas como na esfera pública.

Outra tendência na relação da sociedade com os dados científicos é o aumento de atividades de ciência cidadã (do inglês “citizen science”). A ciência cidadã busca engajar voluntários não especializados em atividades relacionadas à ciência. Um exemplo interessante é o WikiAves<sup>3</sup>, plataforma digital na qual observadores de aves divulgam dados sob a forma de imagens, sons, textos e outros. Redes como essa possibilitam a interação entre especialistas e interessados e colaboram com a compreensão de fenômenos associados à biodiversidade e com a conservação de espécies e ecossistemas.

Os novos interesses e padrões de interação com o conhecimento científico estão em consonância com o movimento mundial em prol da ciência aberta, ou “open science”, que busca tornar dados cientí-

---

<sup>3</sup> O portal WikiAves é direcionado à comunidade de observadores de aves, os quais podem contribuir com conteúdo no endereço <http://www.wikiaves.com.br/>.

ficos acessíveis a todos os níveis da uma sociedade, facilitando a comunicação sobre ciência entre profissionais e amadores em benefício de todos. O movimento é baseado no princípio de que dados acessíveis proporcionam escolhas infor-

mas sobre ações tomadas no mundo. Nessa perspectiva, a existência de dados abertos sobre biodiversidade permite a tomada de decisões com base em fatos e em conhecimentos e deve ser estimulada no Brasil.

---

## SUGESTÕES DE LEITURA

Um panorama dos dados existentes sobre a biodiversidade brasileira e as principais lacunas do conhecimento é apresentado nos dois volumes do livro “Avaliação do Estado do Conhecimento da Biodiversidade Brasileira”, de Thomas Michael Lewinsohn, publicado e disponibilizado na internet pelo Ministério do Meio Ambiente (2006). Para saber mais sobre o repositório de dados do PPBio, com ênfase em dados de vegetação, consulte o artigo “The Brazilian Program for Biodiversity Research (PPBio) Information System”, escrito por Flávia Pezzini e colaboradores e disponível em <http://bit.ly/1QBerT4>. Há diversos portais para consulta: “Biodiversity Information Standards” ([www.tdwg.org/](http://www.tdwg.org/)), traz informações sobre padrões adotados nacional e internacionalmente para dados de biodiversidade; DataOne ([www.dataone.org/](http://www.dataone.org/)) apresenta dados de organizações de todo o mundo, informações sobre boas práticas de gestão de dados, além de eventos e webinars. Dados sobre espécies, inclusive de vouchers coletados no Brasil e depositados como material testemunho em coleções brasileiras estão disponíveis no Sistema de Informação sobre a Biodiversidade Brasileira – <http://www.sibbr.gov.br/>, o qual está vinculado à iniciativa internacional Global Biodiversity Information Facility - <http://www.gbif.org/>. Altmetric ([www.altmetric.com](http://www.altmetric.com)) apresenta métricas alternativas e é voltado a instituições, financiadores, editoras e pesquisadores. A Rede pelo Conhecimento Livre (<http://br.okfn.org/>) busca tornar a relação entre governo e sociedade mais transparente. Sobre ciência cidadã, sugerimos o portal [www.citizenscience.org](http://www.citizenscience.org). Sobre curadoria de dados em repositórios institucionais, sugerimos a leitura do artigo “The Data Curation Continuum: managing data objects in institutional repositories”, de Andrew Treolar e colaboradores, disponível em <http://bit.ly/1ThCJan>.