

PRESS NOTICE

For immediate release

Using 'big data' to help conserve life on Earth

Save the date: Wednesday, 9 October

Launch of overhauled portal from Global Biodiversity Information Facility; Science Symposium to showcase use of data in tackling conservation, climate responses

A meeting in Berlin next month witnesses the unveiling of a new tool making it much easier for researchers and policy makers to discover and use vast quantities of data about the diversity of life on Earth.

The Global Biodiversity Information Facility (GBIF) will use the 20th meeting of its Governing Board to launch a new version of its global web portal, providing greatly enhanced access to the world's largest database of documented evidence for the distribution of species across the planet.

The meeting will also demonstrate how researchers in many countries have used the data mobilized through GBIF's network of data publishers and Participants to tackle questions critical to reducing the loss of biodiversity and the services it provides to people.

The enhanced portal marks a new phase in GBIF's mission to support governments, scientists and wider society by:

- developing the tools, skills and human networks to build the best possible picture of where every species has been found, bringing together data from natural history collections, field researchers, remote-sensing tools, citizen scientists and other sources;
- offering national agencies a cost-effective infrastructure to share and exchange fundamental biodiversity data, supporting a broad range of policy decisions including control of pests and diseases, promoting food security, priorities for wildlife protection and responding to climate change;
- providing a critical information component to global efforts to reduce the pressures leading to biodiversity loss, for example through the Aichi Biodiversity Targets and the new Intergovernmental Science-Policy Interface for Biodiversity and Ecosystem Services (IPBES)
- supporting a growing body of data-driven scientific publications from around the world, with more than 20 peer-reviewed publications per month now citing use of GBIF-mediated data.

The portal launch is open to the press and public, and will be streamed live via the Internet. It is accompanied by presentations of how GBIF's partners in Colombia and Japan are providing free and open access to information about biodiversity in their own countries, using open-source tools and resources developed by the global network.

Later on the same day, the annual GBIF Science Symposium, also to be streamed live, will showcase key uses of the data made available through the network, including the identification of priority areas for plant conservation in South America; and an investigation of how tropical forests are responding to climate change.

All are welcome to attend the launch and Science Symposium - please register free at ...
<http://gb20.gbif.org/GB20/GB20reg/freeReg0>

For those unable to attend the meeting in Berlin, the events will be available for live streaming and subsequent video access from <https://new.livestream.com/gbif/>.

For more information, and to arrange interviews before or during the event, please contact:

Gesine Steiner
Press Officer
Museum für Naturkunde
Berlin, Germany
Tel. +49(0)30 2093 8917
e-mail: gesine.steiner@mfn-berlin.de

Sampreethi Aipanjiguly
Communications Officer
GBIF Secretariat
Copenhagen, Denmark
Tel. +45 35321497
email: saipanjiguly@gbif.org

The GBIF data portal – what's new

- **Greater visibility for national GBIF Participants**
 - *Before:* Information on the data and news from the countries involved in GBIF was dispersed across different platforms and often difficult to interpret.
 - *New portal:* Country pages for each GBIF Participant present clear and comprehensive information about data published from and relating to the country, contact details and news including research uses of data.
- **Greater visibility for institutions sharing data**
 - *Before:* The large and growing community of data publishers who provide the underlying data served through GBIF were not sufficiently visible to users of the portal.
 - *New portal:* Enables strong branding of the contributing data publishers, including display of logos and richer attribution of all those involved in curating and sharing data.
- **Speed of data indexing and availability***
 - *Before:* Publishers sharing data through GBIF often had to wait months before their changes or additions were visible on the portal or available through web services.
 - *New portal:* When data modifications are made using GBIF's free publishing tools, those changes will be harvested in real time thanks to a scalable data crawling infrastructure, and will be visible on the portal and via web services within hours or minutes. Such responsiveness of the system should encourage data quality issues to be corrected quickly, and paves the way for services alerting users to new data meeting specified criteria.
- **Improved search and mapping**
 - *Before:* Finding and visualizing data through the GBIF data portal was limited, making it difficult to understand data characteristics (size, biases, coverage, etc) before downloading.
 - *New portal:* Species occurrence records, species-level information and taxonomy, metadata about datasets and publishing institutions are easily searchable through a wide range of filters. Improved data metrics (multi-dimensional) and high-resolution maps provide better means for users to assess content before downloading, and to report on data quality issues.
- **Removal of download limits and access to verbatim data**
 - *Before:* Users wishing to download more than 250,000 records could not do so using the portal, but had to request special assistance from GBIF and sign letters of agreement regarding access to verbatim data as supplied by data publishers.
 - *On new portal:* Thanks to greatly improved processing capacity, users can freely download datasets of any quantity, and also have access to verbatim data through the portal. Improving the transparency of processing by exposing the content at each stage allows those working with data quality to understand issues better.

- **Making it easier for others to build on the GBIF portal**
 - o *Before:* Web services deployed with the data portal in 2007 have degraded in responsiveness with increased data volumes. The formats used have also become somewhat dated.
 - o *New portal:* The portal is itself built on a new open web services platform. By making use of modern 'Big Data' technologies for processing and search, the web services are designed to scale with increasing data volumes and access, positioning the portal as a robust component to build upon.
- **Linkages to data use cases, news and associated tools and resources**
 - o *Before:* GBIF's data, communication and networking activities were scattered across different platforms, making it difficult to communicate GBIF's combined functions and services.
 - o *On new portal:* Data search and communication are combined into a single platform, enabling users to discover and access examples of research uses, tools, methodologies and national activities – thus helping to demonstrate GBIF's value to all stakeholders.

*This feature is implemented, but remains under testing and will be visible on the new portal before the end of 2013

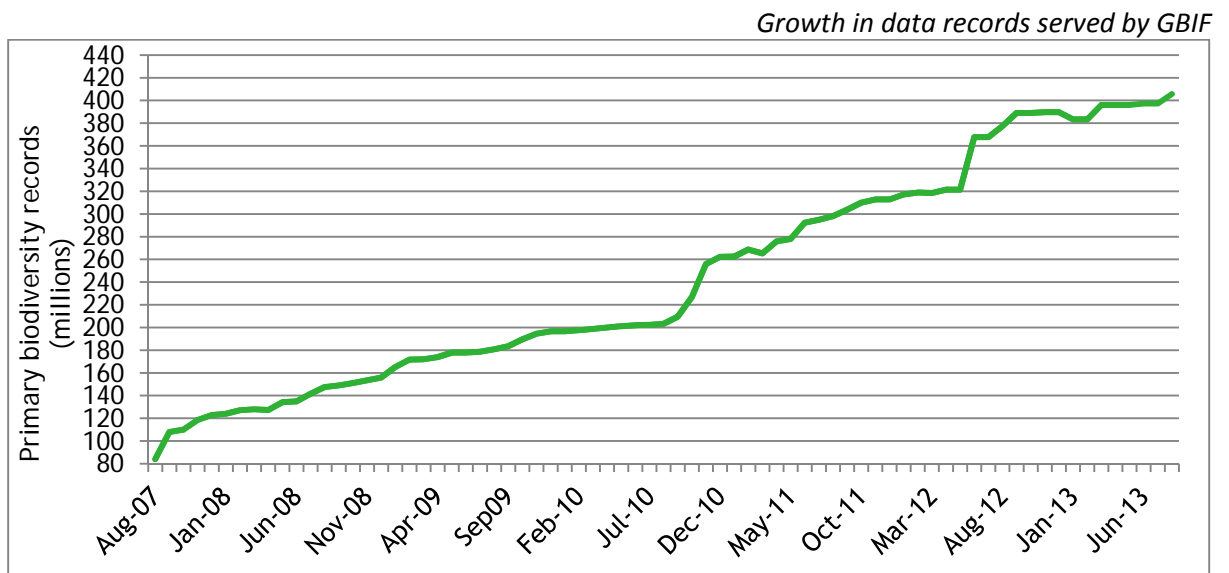
BACKGROUND INFORMATION

The Global Biodiversity Information Facility

- The Global Biodiversity Information Facility (GBIF) is an intergovernmental scientific infrastructure, set up by governments in 2001, to encourage and facilitate **free and open access to biodiversity data**, via the Internet.
- It enables anyone, anywhere to discover and use information about all types of life on Earth, gathered over centuries and held in many different formats. The data mobilized by GBIF's network of countries, institutions and individuals support science, decisions and public awareness about biodiversity.
- GBIF operates through a network of national and thematic 'nodes', coordinating mobilization and use of biodiversity data within its Participant countries and organizations. This includes collaborating to share tools, skills and experience to help set up national biodiversity portals to serve the needs of researchers, decision makers and the public in Participant countries.

The GBIF data portal – a window on biodiversity

- The GBIF data portal offers a single online access point to over **400 million biodiversity records** from more than **10,000 datasets** published by more than **500 institutions**, ranging from museum specimens collected from the earliest days of natural history exploration, to current observations by 'citizen scientists' and monitoring from research expeditions.
- Since its launch in 2007, the volume of data made accessible through the GBIF data portal and associated web services has continued to grow.



GBIF-mediated data – a tool for science and society

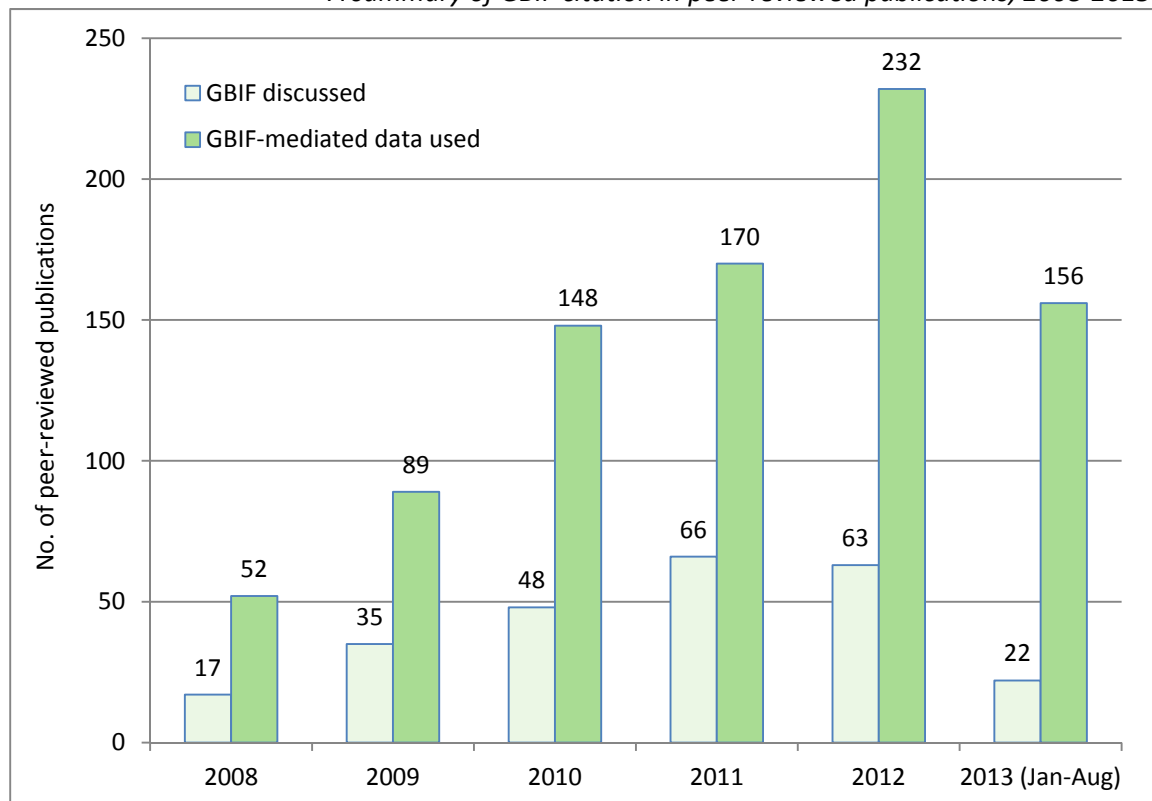
Data discovered accessed through GBIF are being used in many areas of scientific research and decision making. These uses support international policies and responses to key targets for ending the loss of biodiversity.

GBIF-mediated data have been used in more than 800 peer-reviewed scientific studies, for example to:

- Model the potential spread of known invasive alien species, under current and future environmental conditions
- Predict the consequences of climate change for biodiversity, from single species to global impacts
- Research agriculture and food security, including crop wild relatives and pest control
- Suggest priorities for conservation of threatened species and protected areas
- Explore links between biodiversity and human health, including modelling of disease vectors

Examples of research using data accessed via GBIF can be found in the annual GBIF Science Review (http://www.gbif.org/orc/?doc_id=5287) and bimonthly GBits Science Supplement (<http://www.gbif.org/communications/resources/newsletters/>)

A summary of GBIF citation in peer-reviewed publications, 2008-2013



Recent examples of scientific studies using GBIF-accessed data

1. Protecting plant species diversity

A study using more than half a million plant occurrence records downloaded via GBIF, covering over 16,000 species, developed a network of 'virtual parks' to assess how well biodiversity is currently protected in South America.

The map of 'virtual parks' comprised areas where a large number of threatened and locally-unique species occurred within a 100km area. These were found to correspond well with real, existing protected areas – by prioritizing 24 protected areas, the researchers argue, up to 70 per cent of South American plant diversity will be conserved.

They also reveal gaps in existing coverage, suggesting new measures to conserve 200 plant species not currently included in any protected area. Critical areas to monitor, expand and strengthen are mainly located in the Ecuadorian and Colombian Andes, southern Paraguay, the Guyana shield, southern Brazil, and Bolivia.

The lead author of the study – Julian Ramirez-Villegas - said the analysis would not have been possible without the ability to access occurrence data through GBIF.

Julian Ramirez-Villegas is scheduled to speak at the GBIF Science Symposium on 9 October 2013.

Citation: Ramirez-Villegas, J., Jarvis, A. & Touval, J., 2012. Analysis of threats to South American flora and its implications for conservation. *Journal for Nature Conservation*. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1617138112000830>

2. Global study of climate impact on species

Research published in the journal *Nature Climate Change* used data on 50,000 common plants and animals, accessed through GBIF, to forecast major reductions in the climate ranges of common plants and animals around the world.

The study, by researchers in the United Kingdom, Australia and Colombia, found that more than half of the plants and over a third of animal species could lose more than half of their climatic range by 2080 if nothing is done to limit greenhouse gas emissions.

More on this story is at <http://www.gbif.org/communications/news-and-events/showsingle/article/gbif-enables-global-study-of-climate-impact-on-species/>.

The abstract of the research paper is at <http://www.nature.com/nclimate/journal/v3/n7/full/nclimate1887.html>. Subscription is required to access the full peer-reviewed article.

Citation: Warren, R. et al., 2013. Quantifying the benefit of early climate change mitigation in avoiding biodiversity loss. *Nature Clim. Change*, advance online publication. Available at: <dx.doi.org/10.1038/nclimate1887>.

3. Predicting the spread of allergenic invasive plants in Europe

A study by a research team in Germany aimed to predict whether the invasive plant – the common

ragweed (*Ambrosia artemisiifolia* L.) – was likely to shift its range in Europe due to climate change.

Ragweed is a plant native to North America, introduced accidentally to southeastern Europe in the 19th century. It has since become widespread in parts of the continent. Typically growing in urban wasteland and overgrown fields, it is a significant human health risk because many people are allergic to its pollen.

For their study, the researchers tested various models that map the habitats suitable for a species to live in, based on the conditions of temperature and rainfall in the places where it has been observed, and projections for how climate change would affect the locations of suitable habitats (ecological niche models).

To generate the models, the research used GBIF to identify 2,016 records of the ragweed's occurrence in its native North America, and 2,779 records from its invasive range in Europe - together with past climate records and forecasts for climate change based on various scenarios. It compared the results for current suitable habitats with independent data on the regions where ragweed has been reported.

The study found that when only the European occurrences were used, the model produced implausible results. The authors concluded that this was due to sampling bias among the European records available through GBIF when the models were generated (in 2009). While many ragweed records were published from Germany and other parts of northern Europe, few or no records were available from regions where the plant was known to be a problem such as in Italy and Hungary.

However, when the GBIF-mediated data from the plant's native range in North America were used to generate the models, there was a much better match with known occurrences in Europe. On this basis, the researchers predicted that climate change would enable the ragweed to thrive in many more parts of Europe, with potential invasions possible over huge areas including northern France, Germany, the Benelux countries, Czech Republic, Poland, the Baltic States, Belarus and wide parts of Russia.

In view of the high cost of the plant's spread in terms of human health (estimated at €110m per year in Hungary, for example), the authors conclude there is a strong need for control measures to minimize the further spread of common ragweed.

The study has attracted attention from German-language media, including an article in the online version of Die Welt (<http://www.welt.de/gesundheit/article115002629/Ambrosia-verbreitet-ihren-Schrecken-in-Europa.html>), which mentions that the research relied on data accessed through GBIF.

Citation: Cunze, S., Leiblein, M.C. & Tackenberg, O., 2013. Range Expansion of *Ambrosia artemisiifolia* in Europe Is Promoted by Climate Change. *ISRN Ecology*, 2013, pp.1–9. Available at: <http://www.hindawi.com/isrn/ecology/2013/610126/>.

Some figures for Germany

Scientific papers published using data accessed via GBIF

Germany ranked fourth based on the number of scientific publications in 2012, which cited use of GBIF-mediated data. The ranking was based on affiliations of authors, and 21 papers were published last year alone.

In total, 71 scientific papers have been published in peer-reviewed publications by authors affiliated with institutions in Germany, over the last 5 years (2008 to August 2013).

Data records published

About Germany

- Over 2000 datasets containing more than 6 million records.
- 34 countries contribute data about Germany.

From Germany

- 8,400 datasets with over 10.5 million records.
- Germany publishes data covering 243 countries, territories and islands.