

# Frequently asked questions

## General questions

### Why is the NBN proposing to make these changes to the access controls and download functionality?

There are three main “drivers”:

1. **To increase use and data flow.** These proposals would be a significant step forward in making it easier to share information as requested by many data providers and users of the NBN Gateway. This is based on the first of the [NBN data exchange principles](#), which states that “*Biodiversity data should be easily accessible to enable their use for not-for-profit decision-making, education, research and other public-benefit purposes.*” They would also enable effective analysis of data as primary evidence for decision making. We have also been asked by several data providers to set up more flexible access controls that allow access to a dataset to be restricted geographically or taxonomically. This type of filtered access could greatly increase data flows between National Schemes and Societies and Local Record Centres.
2. **To increase clarity.** The current access controls are complex and confusing for both data providers and users. The NBN is aware of instances where incorrect access controls have been applied to datasets, with unintended consequences.
3. **To improve the performance of the Gateway.** As the number of records that the NBN Gateway holds increases, the current access controls are affecting performance. In the long term these issues need to be addressed.

### How do these changes fit with the [NBN Strategy for 2010-20](#)?

The Forward to this strategy states that:

*“The challenge is to increase the use of biological records through improving their quality and the flow of data through the system, whilst maintaining and improving the underlying recording and data access infrastructure. This is a significant task and will depend on the members of the NBN maintaining their resolve to deliver our vision of: “Making all biological records freely and easily available to everyone”*

The NBN Gateway is a very important tool in meeting this challenge, but the current functionality needs to be improved for the reasons given in response to the previous question.

### How do the access controls work at present?

The blurring control provides access to gridded data at a set resolution. Currently these resolutions are 10km, 2km, 1km, and 100m. Each record is blurred to the best access resolution a user has. The sensitive privilege allows a user to access records that have been marked as sensitive, the recorder privilege controls access to recorder and determiner names, and the attributes privilege controls access to the record attributes.

The download privilege is different and does not fit in with the others. It controls how the data may be accessed rather than what data may be accessed. It controls access to the raw records on the Gateway rather than any statistical summaries such as a count of the number of records.

When a user has access via being a member of an organisation as well as personal access rights, the aggregation of controls is via an additive mechanism. This can be non-intuitive. For example RSPB gives Fred Smith access to sensitive records but restricts access to 10 km square resolution. However Fred Smith is a member of an organisation that RSPB has granted access at 100m resolution but not sensitive access. At the moment with the additive nature of the access controls Fred Smith would end up with having sensitive access at 100m resolution which was not the RSPB's intention.

There is a useful guide to the current access controls on the [NBN website](#)

### **Why do we need access controls at all?**

Many of our data providers require the ability to limit access to their data. This can be for conservation purposes, protecting nesting sites or locations of particularly sensitive species from disturbance. It can also be for data protection and intellectual property protection. Controls are sometimes needed to protect the business model of data providers. The 7<sup>th</sup> NBN data exchange principle, is relevant here, specifically: *“Biodiversity data suppliers should try to arrange resourcing of information provision so that charges for not-for-profit uses are minimal and charges for commercial uses are realistic but do not prevent the use of biodiversity data. Biodiversity data users should expect to contribute to sustaining the provision of biodiversity data through contributing either in kind or financially to the collection, collation and management of biodiversity data, or at the point of use”*

### **Do these proposals only affect access and download of species records?**

Yes. Access to habitat data or site boundary layers on the Gateway is not affected by these proposals. Habitat and site spatial layers are not subject to the same complex access controls. They are currently only available to view as backdrops to the species records and are not downloadable.

### **Why is the NBN proposing to remove the separation of download and view access?**

In short, to improve clarity, increase use and improve technical performance.

The current separation confuses both users and staff alike! It is a widely held assumption that ‘What You See Is What You Get’, and when a user requests data behind a map and only gets a partial return of the data it causes a lot of confusion.

At present, the NBN Gateway is typically only providing a basic data discovery service, despite the fact that the new interactive map tool is excellent for exploring and visualising species distributions. This is partly because a significant proportion of records cannot be downloaded – they can only be viewed through the Gateway, and even then the data have been coarsened in many cases. This change would represent a significant shift in change the emphasis of the NBN Gateway from a visualisation tool to a portal that provides access to data that can be used for a wide variety of purposes. The visualisation functionality would be a means to an end and not an end in itself. In effect the Gateway would be increasingly used as a method of producing a “custom download” based on a set of filters.

The removal of the separate “view” access control may increase the volume of UK data contributing to the [GBIF portal](#), but this will depend on how data providers set public access.

Removal of this control will also give a performance boost to the NBN Gateway and web services. This is because the proposed simplification of the Access Controls in general (including dropping the separate “view” control) means that less data processing is required in returning results to a user request.

### **When will these changes be implemented?**

This will depend on the response to this consultation (which ends on 30 November 2011). We might decide to phase in some aspects of these proposals ahead of others, eg to work on proposals that have been positively received, provided that they can be implemented in isolation. Changes might begin to appear from April 2012 onwards.

### **What will happen to existing data access permissions?**

Existing permissions will be transferred to the new system wherever possible, to avoid data users having to go 'back to square one' and re-apply for access. However, in some cases it will not be possible to transfer access levels, for example: If a user currently has full access but with no permission to see recorder names, this access level cannot be transferred to the new system as the user would then be able to view and download recorder names. However, if the data provider decided to reload their data onto the Gateway without recorder names, the existing access level could be retained. On transferring to the new access control system, no one will be automatically receive increased access privileges. In some cases, users might automatically receive decreased privileges and will need to re-apply to the data provider. The good news is that the application process will be much simpler and quicker under the proposed new system!

### **What impact would these proposals have on existing web services?**

There are many different types of NBN web-services, some currently requiring view access (eg the grid map service) and others requiring download access. Some of these services only deliver publicly accessible data, others pick up individual or organisational access privileges. If and when the proposed changes to the access controls were to be implemented all data available through the NBN web services will continue to be subject to the same access controls used throughout the NBN Gateway, in other words the access controls apply equally to all methods of dissemination of data on the NBN Gateway. This means that any "public web-service" will NOT include any sensitive records, recorder names, determiner names or other attribute fields. If you use the web services as a registered user (using your registered username) and have enhanced access to the relevant data you will be able to access all of the data (including attributes) at full resolution (at a max resolution of the records within the dataset). Access to sensitive data will still be granted (or refused) by the data provider on request in addition to enhanced access.

Any changes to the levels of access set by the Data Administrators will however have an effect on web services – for example if public access to a dataset is withdrawn, any web service that retrieves and displays public data would be affected. We hope that most of the existing privileges will be transferred to the revised NBN Gateway access controls and that web services delivery will generally not be affected. There may be benefits in greater clarity of access controls, and in technical performance ie maps and data may be retrieved from the NBN Gateway more quickly.

It is not proposed that there will be a maximum size on web service calls (as suggested for download). It is unlikely that web service developers will need to update any of their code or parameters as a result of the proposed changes. However, because we are not aware of all instances of web service use we cannot guarantee this – please contact the NBN team if you have any specific queries).

## **Will the public be able to see more data or fewer data as a result of the proposed changes?**

This will depend on how data providers respond to these proposals, especially the plan to drop the “view only” setting. It is possible that some data providers would decide to blur their publicly accessible data to 10km square resolution. If we judge that these changes would result in a significant drop in current levels of public access, we would have to take this into careful consideration in deciding how to implement these proposals.

Setting zero public access to a dataset should only be done in exceptional circumstances eg where every record in the dataset is sensitive or as an interim measure, for example while a dataset is undergoing validation.

## **How, specifically will sensitive records be protected?**

Although the streamlining of access controls would help to add clarity to the treatment of sensitive records, we are not proposing any specific changes to the handling of sensitive records. The Data Administrators will continue to highlight individual records as being “sensitive”. Sensitive records will never be made available to the public, and are only available if the Data Administrator explicitly grants “sensitive data” access on a dataset to a specific user or organisation. In effect sensitive records remain totally confidential, available only on a “need to know” basis determined by the Data Administrator.

## **Can you clarify what is meant by a “sensitive record”?**

Sensitive species records are flagged by the Data Administrator according to their judgement of sensitivity. Sensitivity is judged on the basis of the potential risk of damage to the environment. Access to these records will be restricted in accordance with the NBN exchange principle no 2 “*Making biodiversity data available should reduce the risk of damage to the environment. If it is likely to have the opposite effect, availability may need to be controlled*”. This typically applies to rare and protected species. There may be some benefit in agreeing a standard list of “sensitive” species nationally (some progress has already been made on this), but this could never be rigidly applied – for example there may be variation in policy in different parts of the UK or a species may be sensitive at some times of the year (eg the breeding season) but not at others. Please note that a record should not be flagged as sensitive to protect intellectual property rights or solely because of the Data Protection Act.

## **Questions from current and potential data providers**

### **What advantages do these proposals offer to the data provider?**

The most advantageous aspect of these proposals for many data providers is likely to be the ability to provide filtered access using geographical or taxonomic filters. For example it will be possible for a dataset administrator to set full access to members of a specific organisation for all molluscs in Derbyshire. We hope that this will be a significant step forward in making it easier to share information, and this has been requested by many data providers.

The improved clarity of the access controls will also make data sharing more straightforward.

The improved feedback from data downloads (in the form of monthly emails supplied by the NBN) will be invaluable to many data providers, in showing the extent of the benefits arising from their services.

**I work in a Local Record Centre – do these proposals offer any advantages to the LRCs?**

Local Record Centres may benefit considerably, because the geographical filtering may help them to obtain specific download access to more of the National Schemes and Societies datasets (without having to process them). They can use these data (with permission) in their role as “biodiversity data brokers” in servicing requests specifically relating to their geographic area, eg in Environmental Impact Assessments. The Gateway may become a vital data collation tool for the LRCs in this respect.

**I work for a Local Record Centre and want to share a series of records with a National Scheme. Will the access controls enable me to do this?**

Yes. If the National Scheme is registered as an organisation on the NBN Gateway as the Data Administrator you can grant access to specific records. All users who are members of this organisation will gain access to the data.

**The geographical filtering looks very useful but how will it work?**

The geographical filtering requirement most commonly requested by data providers is the ability to filter according to vice County. We would like to implement this as soon as possible, and we will also look at other ways of filtering geographically (suggestions would be welcome as part of this consultation process). If a geographical data layer is already on the Gateway (eg the SSSI or LRC boundary layers) this is likely to be easier to implement as a geographical filter. The interface would however require careful design so as to avoid confusion but would effectively enable you to select a specific boundary or boundaries (eg a vice county) with which to cut the datasets both in your search as a user or when providing access to your data as a data provider. This means that the data you provide access to is specific to the area requested rather than having to provide access to whole datasets which may include national distributions.

**Will it be possible to include sensitive records when granting access to a request for data for a specific geographic area?**

Yes. You can grant access to one or more datasets for a particular area or areas (eg vice county) for either just sensitive records or just non-sensitive records or both sensitive and non-sensitive records. Effectively sensitive/non-sensitive is a mandatory filter which you must set in the same way you must set to which dataset/s you are granting access before being able to proceed with the further detail of the access request.

**Will I be able to use more than one of the proposed filters at a time?**

Yes. The filters will be additive meaning that you can select as many of the filters as you require. This includes the same filter several times (ie selecting more than one geographic area).

**I am concerned that the enhanced ability for users to download data will not be adequately controlled. How will this be regulated?**

We will monitor the use of access controls to make sure there are no “loopholes”. The streamlined controls will make this task easier. Only registered users will be able to use the download functionality and what they are able to download will continue to reflect the access

that they have been granted. If no enhanced access has been applied for, or granted, then the registered user will only be able to download data at the level set for public access. Unregistered users cannot download data from the NBN Gateway.

Users of data downloaded from the Gateway have a responsibility to abide by the general NBN Gateway terms and conditions and any specific additional requirements set by the dataset administrator. We will monitor data downloads closely and provide feedback to dataset administrators. Users will be asked to provide details of how they intend to use the downloaded information and to specifically tick a box to acknowledge that they will abide by the [terms and conditions of use](#).

The NBN Gateway Terms and Conditions contain several clauses designed to prevent misuse of information eg

*2.6 You may not make any financial profit from use of the material, data and/or information on this website or from any products you derive without first obtaining written permission from the relevant data provider.*

*2.7 You may not republish wholesale the material, data and/or information made available to you, or exploit it for commercial or academic research purposes without first obtaining written permission from the relevant data provider [the contact details of which are provided within the metadata of the relevant dataset].*

We will provide a mechanism for members of the wider NBN community to notify us of potential breaches of the T&Cs, and we will consult with the relevant data providers on what measures should subsequently be taken.

**I want to give access for 6 months only to an MSc Student, but only to records for a given taxonomic group in a certain vice-County. Will this be possible under the new access controls?**

Yes. As long as the boundary you want to use is on the NBN Gateway (and in this case the Vice-Counties are) then it can be used to provide access to a subset of a dataset. The ability to set a time limit for data access is also part of our plans.

**I want to grant someone access to a single taxonomic group within my dataset. Will this be possible?**

Some users would seek access privileges to formal taxonomic groupings, whereas others need data to be based on informal groupings that do not correspond exactly to the “official” taxonomy. Examples of the latter include water-beetles, trees, seabirds, waterfowl or butterflies.

The Natural History Museum is currently working on completing the formal taxonomic groupings and hierarchies and we hope that these will be incorporated into the NBN species dictionary by the end of March 2012. This will enable filtering on all formal taxonomic groups, for example families of diptera (two-winged flies).

The informal groupings are not currently set up on the NBN Gateway – we will work with the NBN community to define these, and incorporate these within the NBN Gateway functionality. The timescale for this has not yet been set.

**As a Data Administrator will I be spending all of my time servicing access requests and fine tuning access controls?**

The enhanced download functionality is likely to lead to an increase in the number of access requests but the scale of this will depend on how the Data Administrators apply the access controls. A conservative approach to assigning privileges will generally mean a higher volume of requests for full access.

The controls for granting and revoking access will however be much easier to use than the current mechanism. In addition, you will still be able to quickly grant access to all records within a dataset. With the interface improvements we hope you will be spending less time administering datasets, than you do fighting the current system. We will try to develop an interface that helps Data Administrators in managing a large portfolio of datasets (suggestions for this interface would be welcome from such data providers).

During this consultation process we will “mock up” a series of screenshots to show how the new access controls interface might appear to the Data Administrator.

**Will download access be restricted only to registered users?**

Yes – before downloading, users will have to log-in, which of course means that they will need to be registered. As at present, anyone can apply to be an NBN Gateway registered user. All registered users will have download access to the data that have been granted public access by the data provider.

**Will I be able to see who has downloaded my data and what they will be using it for?**

Yes. Each download of data will be logged and included in a new set of usage statistics and information pages tailored to each data administrator and will also be sent to data administrators as part of a redesigned monthly email. Prior to downloading data, users will be asked to provide details of how they intend to use the downloaded information, the use category (research, casework etc) and to specifically tick a box to acknowledge that they will abide by the terms and conditions of use.

These usage logs and statistics will be invaluable in providing evidence of use, and could be very useful to data providers in communicating with funders and recorders.

**Would I be able to set access levels for a specified geographic area (eg a Vice County) for some users to 1km, others to 2 km and a third category with 10km access?**

No. An important part of these proposals is that there are just two levels of access: public and full. Allowing intermediate blurring (ie 1km, 2km, 10km) would be excessively complex in terms of the technical implementation (vast number of permutations that would need to be calculated on the fly) and therefore result in prohibitive performance delays.

Therefore, it is not possible to implement the proposed filtering (geographic, taxonomic, date ranges etc) for both granting access and searching the NBN Gateway with the ability to blur records at intermediate levels. The proposed filtering is only possible with the two access levels of public and full.

## Questions from users

### **Will this mean as a logged in user, that I get access to more or fewer records?**

About the same on average, but this will depend on the access set by data providers.

### **Will it be more straightforward for me to download data?**

Yes. New provision to download the data behind the map and to download targeted cuts of data will be developed. We are currently referring to the latter as a “download wizard”.

### **How will the download functionality work?**

There will be two main methods for downloading data on the NBN Gateway.

1. Via the Interactive Mapping Tool (ie clicking a button to download a series of records displayed on the Interactive Mapper). This means the IMT can be used to explore and filter records prior to downloading.
2. By using the “download wizard”.

The download wizard will make it easier to download whole datasets as well as making it possible to restrict the download to the required subsection of data. We anticipate that users will be able to launch the “download wizard” from the homepage of the NBN Gateway.

In order to arrive at a subsection of data for download the wizard will make use of the proposed new filters. As with requesting and granting access to data, these filters will enable you to choose data for download either taxonomically (species and hopefully families) or by geographic area, year range or individual survey. These filters will be additive meaning that you can add any number of each type and in any combination. The geographic boundaries available in the first iteration are likely to be vice county, standard counties and unitary areas and LRC boundaries with country and other boundaries available at a later date. It will also be possible to preview the filtered results using the Interactive Map Tool, prior to proceeding with the download (not available before March 2012). We would also like to provide a means of saving and retrieving a series of filters, but this will also not be available in the first release of the download wizard functionality.

There is likely to be a maximum file size (either in terms of number of records or actual size of the file) for each download. This is likely to be around 15million records and is to retain an acceptable level of performance.

### **Will it be easier for me to apply for better access?**

Yes. There will be a mechanism to allow you to request only the data you want across a series of datasets. The process of applying for download/improved access (where it has not been set by default) is currently very cumbersome. For example, a typical request for a 10km square report will often require contacting more than 20 organisations. Users are often not prepared to undertake this consultation and will often work with a limited subset of the data that might potentially be available to them.

### **This is all very well, but what sort of response will I get from data providers?**

We will help data providers to provide a more efficient service by giving them better tools to administer data access requests. Requests for better access are sometimes more likely to be granted if they are targeted at a specific group of records and if the requests have a clearly defined purpose. Users are reminded of the 7<sup>th</sup> NBN Data Exchange Principle.

*“Biodiversity data users should expect to contribute to sustaining the provision of biodiversity data through contributing either in kind or financially to the collection, collation and management of biodiversity data, or at the point of use”.*

**I am a member of an organisation. How does this impact on the access rights that have been specifically granted to me as an individual?**

These proposals will clarify to both users and data providers what access individual users have. Unless data providers radically change access grants, you should see no real impact to the data made available to you.