Data Capture, Verification and Access Summary

How will data be captured and verified?

Data is captured using the existing online system (www.npms.org.uk). At the end of each year analysis of data that has been entered using the online data entry system will be fed back to volunteers.

Due to the ability to accurately locate plots using GIS mapping and the entry of species names via currently approved nomenclature, errors should be minimised and all data entered via the NPMS systems will therefore be *validated* against the NBN data standards.

The NPMS benefits from a range of pre-existing data *verification* functions available via Indicia. All species occurrence data entered to the NPMS will be checked against the NBN Record Cleaner rules developed by the BSBI. For example, observations of species beyond their existing geographic range will be flagged. Records of species that are more difficult to identify will also be highlighted. These flags will be used to categorise two sub-sets of the NPMS data.

1) Observations of relatively easily identified species (coded 1 or 2 in the BSBI classification of higher plants according to identification difficulty) which do not occur outside of their known range will be coded as 'auto-verified'.
2) For other observations (i.e. for plants classified with a high level of identification difficulty and/or observations outside of known ranges), NPMS data will be reviewed by expert botanists on a regional basis. The iRecord data verification system will be used for expert review of the NPMS data as it provides an existing, feature-rich interface.

How can data be accessed?

Wide access to **NPMS field data** and derived datasets will be available through a variety of routes. The NPMS website will be the primary route for participants to access their own data, through facilities to download in a variety of formats.

To **encourage and track research use** of the data it is likely that **summary data sets** will be published with DOIs (these are 'Digital Object Identifiers', and enable the usage of specific datasets in scientific research to be tracked) at **the end of each field season**.

Distribution data on the **National Biodiversity Network Gateway** will be published at **full resolution** (in conformance with the data access policy agreed with the project steering group) by the end of the calendar year in which it is collated.

Sensitive information (both species and sites) will be handled appropriately.

Metadata that allow the field data to be analysed alongside **other environmental datasets** by external organisations will also be made available where appropriate.