**JNCC Technical Skills Framework
a strategic approach to training**
Version 1.2
October 2015

JNCC Technical Skills Framework
a strategic approach to training

Introduction

This document provides an initial draft of a framework for technical skills within the organisation. It splits skills into seven strands:

- **Data management, security and audit** – basic familiarity with responsibilities around data management including controlling access to it
- **Spreadsheets (Excel)** – familiarity with the sorts of data manipulation tasks that can be carried out using Excel
- **Databases** – knowledge of the concepts of a database and how to maximise the use of these for storage and use of data
- **Spatial data** – managing, querying and displaying data which has a geographic component
- **Data modelling and "R"** – developing processing chains and models using the statistical package "R"
- **Programming, scripting and application development** – developing scripts both to automate tasks and control interactions with users (though the development of applications)
- **Remote sensing** – knowledge of the sorts of data that can be gathered through remote sensing and the ability to process this to create value added products
- **Statistical knowledge** – familiarity with and understanding of statistical techniques to analyse data including when this can / cannot be applied and confidence in the results

This list is concentrated on the skills that JNCC strategically needs to develop. Generally all scientific staff within JNCC (and any non-scientific staff that are interested) should, as a minimum, have a broad awareness of what these strands are, the sorts of tools that exist within them and what is potentially possible. As part of the general roll-out we will create, and regularly run, an overview presentation which **all relevant / interested staff would be expected to attend** and would also be run for new staff coming into the organisation. Beyond this, line managers working with staff should consider exactly what level of competency within each strand is needed to support both existing work as well as the general development of the individual.

Within each strand three levels of competence are described, the most basic of which is "foundation". The aim is that particularly more technical staff would have at least foundation level experience in many if not all of the strands to ensure that they have sufficient breadth of experience to allow them to judge which tool is likely to be most appropriate for any given task. This is specifically intended to address the risk of an individual developing knowledge of one tool and using this to solve all problems regardless of how effective that tool is for any particular job. The goal is that all technical members of JNCC should have a broad technical base on which more specialist skills are built.

The next level up is "intermediate" and is the most flexible level with many paths an individual can follow as they progress through this depending on their skills, demands of the work etc. Key points:

- Strong linkage to activities within the individual's work plan is critical – Any training but particularly technical training courses should only be undertaken when there are appropriate follow up activities within the individuals work plan. The majority of the real learning in

developing technical skills occurs not in the classroom but in actual application and if there is much of a gap between attending a training course (or other training activities) and applying the learning, much of the benefit of the training can be lost. Additionally, having clear sight of the goals within the work plan will really help to identify the most appropriate course, if this is needed.

- Maximise use of other mechanisms for learning – there are a variety of other mechanisms available for developing technical skills including books, online tutorial / help pages, forums and particularly exchanging problems/solutions with other internal staff. These sorts of mechanism are usually a much more effective way to develop good technical skills as gradually working something out often exposing other avenues that will be useful later and ensure that the solution is more permanently embedded in memory.
- Development of specific skills - In certain situations an individual may develop a very specific skill in the intermediate level. This could be where there is a significant strand of repetitive work that needs to be run once the basic skills have been developed. This is perfectly acceptable though there would be merit in ensuring that the approach is endorsed by an individual with broader experience to ensure the tool is the most appropriate and not simply using the one which the individual is familiar with.

The highest level is "expert" and individuals reaching this level will often be those who have come through a more technically focused education programme (e.g. a degree / masters in a technical domain). However, it is entirely possible for individuals with the right aptitude to progress to this level in post but would require more focused mentoring as well as attendance on a range of professional courses. It is very important that JNCC maintains, at an organisational level, a critical capability at the advanced level in each of the technical strands though not necessarily within a single team.

Data management, security and audit

Foundation

Individual has a good understanding of the importance of data including adequate descriptions (metadata), version control, audit trails and the principles of licensing. They are also very aware of the main policies around transparency and data security and the appropriate balance of risk around these.

Specifically:

- Knowledge of, and compliance with, the Information Management policy, exercise good judgement, and to safeguard and share information appropriately.
- Familiarity with the importance of metadata and able to create good records for data resources within the JNCC metadata catalogue (TopCat)
- Understand the importance of tracking dataset versions for transparency and repeatability and how this is practically implemented
- Awareness of the main policy and legislative drivers around accessibility of data including Environmental Information Regulations (EIR) and Freedom of Information (FOI).

Courses:

- Civil service learning course (Responsible for information: General User)
- Introduction to creating metadata and using TopCat (internal)
- Introduction to FOI, EIR etc (part of induction training)

Intermediate

Good knowledge of what constitutes good information management and can spot and assess risks surrounding the security and integrity of information holdings. At the upper end of intermediate this includes managing and developing other staff to undertake responsibility for individual sources.

Specifically:

- Good general technical knowledge around managing data including metadata, version control etc.
- Very familiar with JNCC Information management and security policies
- Able to assess security risks around sources, classify them and establish appropriate controls around their accessibility and use
- Awareness of issues around Intellectual Property and data licensing and the importance of clarity when generating, using and publishing data
- Familiar with, and experience of, creating Open Data Institute Certificates for datasets
- Develop and oversee staff to ensure that they have the necessary knowledge and are implementing appropriate measures to maintain and protect data and information they are responsible for

Courses:

- Civil service learning course (Responsible for Information – Information Asset Owner (IAO))
- Publishing metadata to data.gov.uk

- Open Data Institute certification

Expert/Advanced

Able to assess the overall organisational policies and practice and ensure these are appropriate and effective. Individuals are able to assess more significant risks to the organisation and identify appropriate action following incidents / near misses.

Specifically

- Able to identify and assess organisational risks surrounding data security including the potential impact of new technologies
- Manage and develop the internal policies and processes which facilitate good information management and organisational compliance with internal information management policies.

Courses:

- Civil service learning course (Responsible for Information - Senior Information Risk Owner (SIRO))

Spreadsheets (Excel)

Foundation

Overview of the core Excel functions for manipulating data, the basic strengths of using a spreadsheet and when a more robust solution (i.e. a database) ought to be implemented.

Specifically:

- Overview of the broad capabilities of a spreadsheet and able to create and manipulate data within a sheet (insert, delete etc)
- Perform simple sorts, filters etc to refine the data displayed
- Able to create simple numerical and textual formulas
- Role of spreadsheets – when to use and not to use

Courses:

- Introduction to Excel (internal)

Intermediate

User is very familiar with the full range of Excel functions and create and edit complex formulas. Able to control the values entered in a particular cell and cross reference lookup values. Also able to link to database sources and manipulate the display of data (e.g. through pivot tables). Understanding of how VBA can be used within Excel to create new functions.

Specifically:

- Experience of developing and using simple functions in Excel VBA
- Able to control data content through the use of lookup tables etc
- Able to visualise and manipulate data from a database source

Courses:

- Intermediate Excel (internal)

Databases

Foundation

User has a basic knowledge of what a database is, why and when it should be used and an understanding of basic querying. Foundation level should be run using MS Access. All technical users within JNCC should have this level of competence.

Specifically:

- Understand the basic concepts and benefits of storing data in a database as opposed to a spreadsheet
- Controlling contents of a field
- Understanding of entities (or tables) and when and why separate entities should be used
- Able to query a table using a simple query wizard
- Knowledge that multiple tables can be joined in a query
- Knowledge that there is an underlying language (SQL) for actually extracting the data

Courses:

- Basic introduction to databases (internal)

Intermediate

User is very familiar with database concepts, and can design schemas and is very familiar with querying complex structures (including directly through SQL). Generally the majority of the work and analysis is being done in corporate databases (particularly PostGres).

Specifically:

- Good understanding of database design theory and capable of designing multi-entity databases
- Familiar with the full range of SQL constructs and can competently create and edit statements from scratch
- Good understanding of the different join types including when they should be used
- Experience of using corporate database systems (particularly PostGres) through the management console

Expert/Advanced

User is capable of managing large corporate databases, including security and user access, tuning database performance, backing up and restoring etc

Specifically:

- The ability to tune indexes based on output from the database engine tuning adviser.
- Locate slow running queries and diagnose performance issues
- Create and run disaster recovery plans
- Operate replication mechanisms to synchronize data across multiple systems

Spatial data

Foundation

The individual has a general awareness of what spatial data is, how to visualise it through a GIS and the sorts of analysis that is possible with it.

Specifically:

- Awareness of the different types of spatial data (points, lines polygons) and raster data
- Capable of loading data into QGIS and altering display
- Able to perform simple spatial queries through QGIS including overlaps, buffering, clipping and dissolving

Courses:

- GIS247 Guide to QGIS Part 1 & 2 (available September 2015)

Intermediate

The individual is able to create publication ready maps in QGIS including pulling data from a range of separate data sources. They are very familiar with storage and use of data within spatial databases (particularly PostGIS) and able to design, create, edit and query these both from the database interface itself and from GIS software.

Specifically:

- Very familiar with QGIS for interacting with spatial database and file based sources and able to generate publication level mapped outputs.
- Able to create and maintain spatial sources within PostGIS and confidently move data into and out of this platform
- Familiar with the concepts of Web Map Services (WMS) and Web Feature Services (WFS) for simple interactions with the spatial data and able to access and use these from QGIS and create simple web based maps using them
- Able to build spatial queries directly in PostGIS
- Broad overview of the capabilities of FME and able to use this for a broad range of data transformations

Expert/Advanced

The individual is able to build and maintain large and complex spatial repositories including tuning / scaling, backing up and restoring, maintaining security and troubleshooting problems that occur.

Specifically:

- Control user access to a spatial data repository
- Backup a repository and carry out a full restore or specific elements
- Troubleshoot problems with performance and tune database to address issues
- Scale the repository through adding additional hardware
- Troubleshoot problems encountered by users of the repository

Data modelling and "R"

Foundation

User has a basic understanding of the broad capabilities of R, including loading data, performing simple manipulations and summaries, using basic statistical functions, creating visual outputs and saving / exporting results. The expectation is that the user has enough of a start to be able to explore the use of the application within a particular area. All technical staff should have covered a course to this level and completing it would be a prerequisite for attending other external courses.

Specifically:

- Overview of different data types and able to import data from CSV file
- Basic data cleaning and transformation including getting rid of incomplete cases, changing units, calculating a new variable from two or more of those imported
- Create simple summary statistics such as mean, SD, median, counts, etc
- Perform linear regression to explore relationship between two variables
- Create simple visual outputs including histograms (to show variation and its distribution of the input variables) and scatter plot with fitted regression line.
- Saving the results including outputting a data table and creating image files from a plot

Courses:

- Basic introduction to using R (internal)

Intermediate

Users operating this level are able make use of R to establish and run more complex tasks including making use of additional packages. Any staff members who have a significant proportion of analysis within their work plans should be at intermediate level.

Specifically:

- Able to interact with external databases (particularly PostGres) to obtain and interact with data
- Able to work with spatial data (including PostGIS) through the GDAL interface
- Familiar with a broad range of transformations
- Good familiarity with R Studio
- Able to independently write and debug R functions including a good understanding of variables
- Able to implement and perform simple adaptations to existing modules and establish efficient and repeatable processing chains to use these.
- Knowledge of code repositories for maintaining code and basic understanding of security risks.
- Output mark down of the analysis that has been done

Courses:

- Accessing data in PostGres (Internal)
- Introduction to R Studio and R programming

- Using R spatially
- JNCC code repository and security (internal)

In addition it is expected that there could be more specialist courses needed to develop knowledge of application within particular areas (e.g. processing VMS data, habitat classification etc)

Expert/Advanced

Expert users are capable of developing new modules to run relatively complex analysis and optimise the performance of these. Users operating at this level will generally have built up a good knowledge of programming concepts in other fields.

Specifically:

- Good understanding of vector arithmetic and vectorization methods for fast manipulation of data
- Good understanding of R scripting and able to create packages including interacting with external libraries (e.g. written in C).
- Familiar with the concepts of Object Orientated programming and able to develop within this framework in R.

Courses:

- Object orientated coding in R

Programming / scripting and application development

Foundation

Users have a broad overview of what is possible using scripting / programming and the basics of what is involved. Foundation level would run using VBA (Excel or Access) or Python. All technical users within JNCC should have this level of competence.

Specifically:

- Broad overview of what programming / scripting is and the sorts of things that are possible. Would cover data inputs and outputs, manipulations and processing, interacting with a user etc
- The concept of variables, numerical and string manipulations and loops
- Experience of creating at least one simple script on a popular platform

Courses

- Introduction to programming (in VBA or Python) (internal)

Intermediate

An intermediate user can competently use programming / scripting technologies on at least one platform. This includes familiarity with all the major programming constructs, use of functions / procedures and ability to interact with a range of data sources (both input and output). Also have a basic understanding of Object Orientated (OO) approach to programming.

Specifically:

- Ability to design and build ephemeral/low usage user interfaces e.g. using MS Access and Access VBA to facilitate data capture and internal use
- Good familiarity with coding in at least one language and familiar with all the major coding constructs
- Understanding of Object Orientated concepts and able to design and implement code using this approach
- Able to develop applications which run over the web

Expert/Advanced

User is a very competent programmer in at least one OO language and able to rapidly switch to new languages. User has a high degree of familiarity with OO design and implementation. Capable of breaking complex builds down into manageable work packages and estimate overall costs.

Specifically:

- Able to design and implement complex applications in a multi-tier environment
- Able to rapidly learn and implement new coding platforms
- Experienced in working within larger development teams

Remote sensing / imagery analysis

Foundation

User has a basic understanding of the sorts of remote capture methods that now exist including differences in the sorts of sensors (and what can be detected), resolution that is possible and a broad overview of the sorts of applications that are possible by combining these. The overview should include both terrestrial and marine, and the full range of platforms including satellites, planes, UAVs, ships etc, as well as an awareness of ways to access both raw and processed data e.g. through online data portals, Google Earth Engine.

Specifically:

- Overview of methods including:
    - the different types of satellite (including pseudo satellites and drones) and the implications of the different orbits , sensor types and capabilities, as well as overall flexibility and availability
    - airborne sensors including aerial imagery, LiDAR and use of drones
    - marine methods including acoustic, multibeam, AUVs and video tows
- Broad overview of the range of things that are discernible using the various methods and the scale at which these operate. In particular determining habitat types, condition of these and ability to efficiently detect change.
- Understanding of how to obtain relevant higher-level remote-sensing products for an area of interest and display/analyse this in a GIS.

Courses:

- Use of ship borne sensors and AUVs in the marine environment (internal)
- Use of satellites and airborne sensors (covering both terrestrial and marine environments) (internal)

Intermediate

Individual capable of establishing processing chains for raw remote sensed data, develop segmentation algorithms to discriminate target features and assess quality of the outputs.

Specifically:

- Understanding of, and ability to apply, image correction processes, such as ortho- and geo-rectification, atmospheric correction.
- Understanding of basic image editing algorithms (e.g. sharpening, blurring, masks) and ability to implement these in a scripting language (Python, R etc).
- Ability to implement basic image segmentation algorithms such as threshold based methods in a scripting language and/or more complex methods in specialist software (e.g. for analysis of raw acoustic data in the marine environment).
- Understanding of QA methods and uncertainty analysis.

Expert/advanced

Able to develop and run more complex algorithms including operating in large scale technical infrastructures to allow heavy weight processing.

Specifically:

- Able to develop more complex segmentation algorithms, including through machine-learning methods.
- Able to process satellite information using hosted computing facilities (e.g. JASMIN/CEMS) and download higher level products locally.

Statistical knowledge

Note that the critical distinction around statistical knowledge as opposed to use of techniques within other tools is the level of understanding behind the concepts being used. So, for example, foundation level does not indicate that the user has derived mean, standard deviation or regression relationships between variables but more that they have a deeper understanding of these concepts.

Foundation

User has a good understanding of basic statistical theory and is able to summarise / visualise datasets, identify relationships between variables and understand what can be inferred from these as well as a having a good understanding of the concept of probability.

Specifically:
- Understanding of the basic properties of data (normal / non-normal distributions)
- Understanding of uncertainty, how this is represented as probability and how multiple probabilities interact with each other.
- Familiar with a range of techniques for identifying relationships between data and what can be inferred from these (linear regression and correlation). This includes regression techniques for a range of data types (e.g. Logistic regression)

Intermediate

User is familiar with the theory and application of a broad range of multivariate statistical techniques.

Specifically:
- Familiar with Analysis of Variance (ANOVA) and Multivariate Analysis of Variance (MANOVA) and applying hypothesis tests on Covariances
- Use of range of other main multivariate techniques including Principal Components Analysis, Correspondence Analysis and Multidimensional Scaling
- Use of Discriminant Analysis for classification / discrimination
- Knowledge of the additional dependencies and complexities which emerge when dealing with spatial data and the techniques for assessing this and integrating into models and hypothesis testing

Expert/advanced

User is familiar with the theory and use of much more advanced modelling techniques capable of application to more complex and dynamic systems. Familiar with how the techniques work and able to assess the limitations of outputs as well as when applying existing techniques to novel areas is likely to generate meaningful insight.

Specifically:
- Range of techniques and theory for modelling dynamic systems including application to techniques such as Lotka Volterra equations.
- Reaction-diffusion equations for modelling invasions
- Stochastic modelling including Markov Chains and various equilibria and how these can be applied to systems such as population modelling