

Biodata Management Mini-Guide

Dan Randow, 18 December 2013. [CC BY 3.0](#) Prepared with support from Horizons Regional Council and funded by TFBIS (project 263: Biodata Management Framework: Phase Two).

About this Mini-Guide

This Mini-Guide is a concise version of the Dataversity Biodata Management Guide that is available at <http://dataversity.org.nz/guide>.

This Mini-Guide will enable you to make a quick overall assessment of the current state of your biodata management and a broad plan to improve it.

You can use this Mini-Guide to assess and plan improvements to a biodata management system of any scope.

How to Use This Guide

To use this Mini-Guide to assess and plan improvements to a biodata management system, take the following steps, using the System Assessment Form.

1. Document the organisation, scope of the system to be assessed, person making the assessment and date of the assessment.
2. Choose a maturity level that best describes the current maturity of data management within your system. Describe how your system meets the criteria for the level of maturity you have chosen. Refer to the following tables to determine the maturity criteria.
 - Maturity with respect to data management Activities
 - Maturity with respect to data management Maturity Factors
3. Choose a target maturity level and a deadline for improving your data management to the target level. Describe the changes needed for the system to meet the reach the criteria for the target maturity level. Again, refer to the tables for the criteria.

Biodata Management Maturity

This Mini-Guide defines mature data management as efficient management of data to ensure its fitness for purpose. It defines the following five levels of maturity.

- **Fragmented** – Data management is ad hoc and inconsistent.
- **Improvised** – Some data management is planned and consistent.
- **Managed** – Data is consistently managed.
- **Automated** – Tools enable consistent and efficient data management.
- **Integrated** – All systems support consistent and efficient data management.

The Mini-Guide provides criteria for each maturity level across five data management Activities (shown on page 2) and with respect to six data management Maturity Factors (shown on page 3). The Mini-Guide also provides (on page 4) a form that can be used to make a quick assessment and improvement plan for a biodata management system.

Maturity with respect to Data Management Activities

Activity	Maturity				
	Fragmented – Data management is ad hoc and inconsistent.	Improvised – Some data management is planned and consistent.	Managed – Data is consistently managed.	Automated – Tools enable consistent and efficient data management.	Integrated – All systems support consistent and efficient data management.
Capture – Data is recorded in the field or at a desk.	Data capture is ad hoc and inconsistent.	Some data capture is planned and consistent.	Data capture is consistently managed.	Tools enable consistent and efficient data capture.	Data capture systems are integrated with other systems.
Ingest – Data is introduced into a primary repository and catalogued.	Ingestion is ad hoc and inconsistent.	Some ingestion is planned and consistent.	Data ingestion is consistently managed.	Tools enable consistent and efficient data ingestion.	Data ingestion systems are integrated with other systems.
Store – Data and metadata are retained for the required term.	Storage is ad hoc and inconsistent.	Some storage is planned and consistent.	Data storage is consistently managed.	Tools enable consistent and efficient data storage.	Data storage systems are integrated with other systems.
Share – Data and metadata are available to the required internal and external people and systems.	Sharing is ad hoc and inconsistent.	Some sharing is planned and consistent.	Data sharing is consistently managed.	Tools enable consistent and efficient data sharing.	Data sharing systems are integrated with other systems.
Analyse – Datasets are combined, compared, summarised and presented.	Analysis is ad hoc and inconsistent.	Some analysis is planned and consistent.	Data analysis is consistently managed.	Tools enable consistent and efficient data analysis.	Data analysis systems are integrated with other systems.

Maturity with respect to Data Management Maturity Factors

	Maturity				
Factor	Fragmented – Data management is ad hoc and inconsistent.	Improvised – Some data management is planned and consistent.	Managed – Data is consistently managed.	Automated – Tools enable consistent and efficient data management.	Integrated – All systems support consistent and efficient data management.
Processes – Processes for managing data are maintained and followed.	Processes are ad hoc and inconsistent.	Some documented processes are followed.	Documented processes are consistently followed. Adoption of processes is measured.	Tools enable consistent adherence to documented processes. Processes are maintained.	Processes and their support by tools are under ongoing review and improvement.
Tools – Tools are used to facilitate data management.	The use of tools is ad hoc and inconsistent.	The use of some (usually generic) tools is planned and consistent.	Generic tools or legacy custom systems are consistently managed.	Recognised off-the-shelf or well-maintained custom tools are used.	Modular, interoperable, cross-platform tools are managed by organisational IT processes.
Formats – Structured and open formats are used.	Unstructured analogue and proprietary digital formats are used.	Structured analogue and proprietary digital formats are used.	Structured open formats are used.	Individual rows are dynamically available with universally unique IDs (UUIDs).	Data uses RDF.
Licensing – Applicable copyright and licence information is recorded and permissions are enforced.	Licence information is seldom recorded and permissions are inconsistently enforced.	Licence information is sometimes recorded and permissions are sometimes enforced.	Licence information is consistently recorded and permissions are consistently enforced.	Licence information is handled and permissions are enforced automatically.	All systems handle licences that vary with aggregation, and over elapsed time.
Reliability – The user has sufficient information to determine how reliable the data is.	Reliability is not explicitly managed.	Reliability is sometimes managed.	Reliability is consistently managed.	Reliability indices are used.	Data depreciation is supported.
Standards – Data and metadata use standard terms.	No data standard is used.	Internal data standards are used.	A data standard, external where one exists, is consistently used and referenced.	References to standards used are machine-readable.	Only well-established standards are used.

Data Management Overall Assessment and Improvement Plan

Use this form for a quick overall assessment and improvement plan for a biodata system. For a comprehensive system assessment and improvement plan visit <http://dataversity.org.nz/guide>.

Organisation:	
System scope:	
Assessed by:	Date:

Overall Current State Assessment

Overall current maturity level of data management:
Description of current overall maturity of data management:

Overall Improvement Plan

Target overall maturity level of data management:	Deadline:
Changes needed to reach target overall maturity level:	