

Digital Data Management

Statistics New Zealand's experiences

Hamish James
Official Statistics Research
and Data Archive Centre

Statistics New Zealand

- New Zealand's national statistical office
- Collects, analyses and disseminates statistics relating to New Zealand's economy, environment and society
- Ensures that information provided to Statistics New Zealand remains confidential, and that it will be used for statistical purposes only.

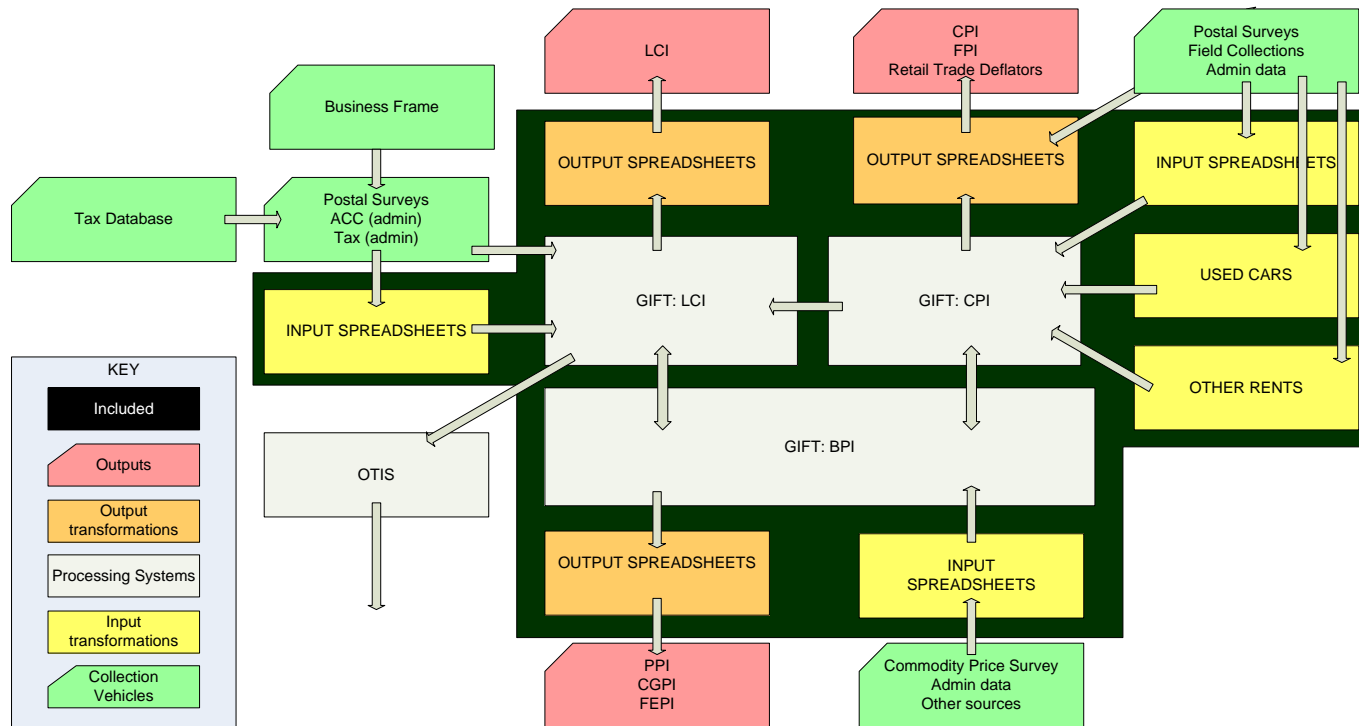
Official Statistics System

- Official statistics are all statistics produced by government departments
- The Official Statistics System (OSS) is all official statistics and the mechanisms that produce them
- Statistics NZ leads the OSS and is the major producer of official statistics

Why data (not record) management?

- Statistics New Zealand's core business revolves around the collection and processing of data
 - Compiling, analysing, abstracting and publishing official statistics
 - People at Stats think 'data', not 'records'
 - Complex data systems

Complex data systems



Corporate directions

- Business model Transformation Strategy (BmTS)
 - Standardising and improving on Statistics NZ business processes and systems
- Top Down Review
 - Improving access to detailed data
- Strategic objective
 - An enduring resource

Common drivers

- More efficient statistical data processing
- Easier discovery of information
- Better comprehension of information
- Easier access to information
- Better provision of data to users
- Legislative compliance

Data management in the BmTS

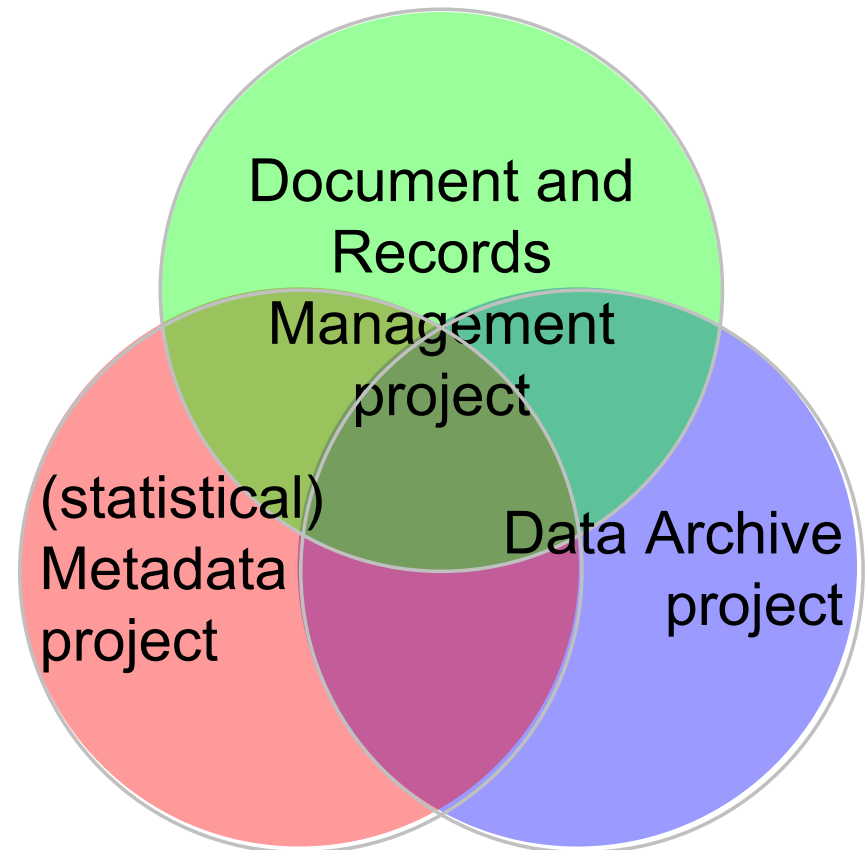
- Employ a disciplined approach to data and information management:
 - Increase the use of administrative data as a primary source
 - Reuse the data already collected to meet a range of different needs
 - Establish metadata as the backbone to the operation of Statistic New Zealand
 - Data is on a need to know basis
 - Access to metadata is open
 - Recognise the life cycle and characteristics of statistical data
 - Data and metadata must be known, accessible, secure and useable

Approaches

Record keeping

Statistical data processing

Digital preservation



Document and Records Management Improvement Project

- Managing corporate information
- Drivers:
 - Legislative compliance
 - Efficient retrieval of information
 - Better sharing of information within Statistics NZ
- Currently: working on introducing a document management system

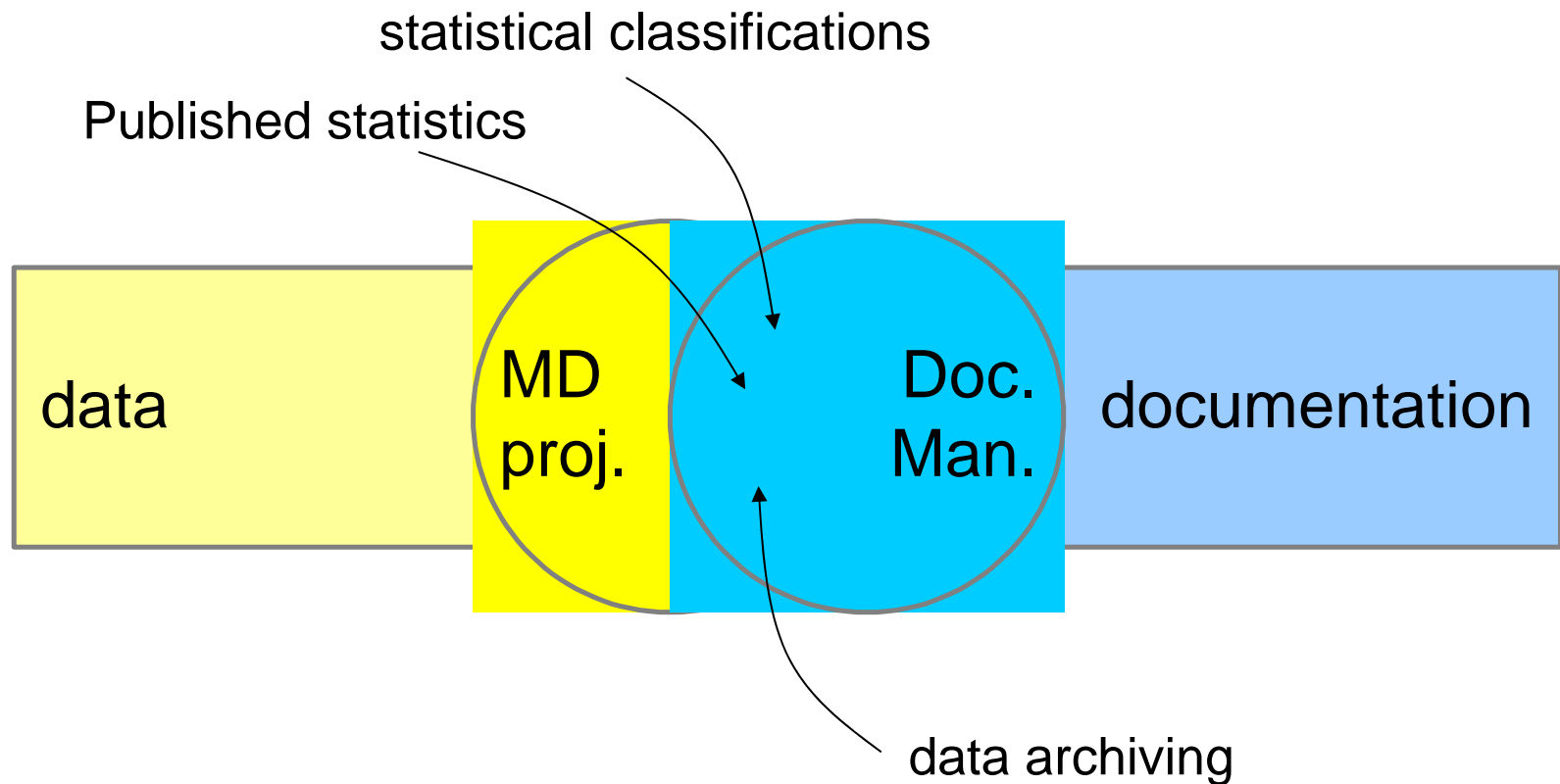
Metadata Project

- Managing statistical data
- Also, Input Data Environment project
- Drivers:
 - Improved ability to interpret and analyse statistical data,
 - Reduce/eliminate the 'redesigning' of processes, business rules and methodologies,
 - Create a framework for managing data from conception to archiving or destruction.
- Currently identifying a statistical metadata model to adopt

Data Archive project

- Managing archived data
- Drivers:
 - Improving researchers' access to unit record data
 - Making data available for reuse, in a usable form
 - Preserving older 'at risk' data
- Currently: appraising data systems and datasets

Data vs documentation



Interesting questions

- Who is responsible for archived documents that describe datasets?
- What is the status of data and metadata held inside unstructured documents?
- Do we need to archive the website at all?
- What is a dataset? What is a document?

Common interests: Culture change

- More generic roles in the organisation means we need good documentation and consistent terminology.
- Shift towards:
 - More consistent practices across all units and offices
 - Greater emphasis on good practice
 - Ensuring compliance

Common interests: Metadata

- Consistent approach to search and discovery across corporate records, operational and archived data
- But, different approaches to metadata
 - Metadata about documents
 - Metadata about data and datasets
 - Metadata about digital objects
 - Statistical Classifications
- Identifying the boundary between structured and unstructured information

Common interests: Information Management

- Working to combine/coordinate:
 - Data custodianship policy
 - Policy for retention and preservation of statistical datasets
 - Principles for managing unstructured information
 - BmTS requirements