Powering Official Statistics at Statistics New Zealand with DDI-L and Colectica

A Case Study
Authors

- Adam Brown
  adam.brown@stats.govt.nz

- Sally Vermaaten
  sally.vermaaten@stats.govt.nz

- Jeremy Iverson
  jeremy@colectica.com

- Dan Smith
  dan@colectica.com
“Turning data into relevant knowledge, efficiently.”

- Ensure New Zealand has the statistical information it needs to grow and prosper
- We do this by:
  - Make sure the right statistics are produced
  - Make sure as many people as possible use the statistics to support informed decision making
Browse for Statistics

Snapshots of New Zealand
- Top statistics
- New Zealand in Profile: 2012
- Te Āhua o Aotearoa: 2012
- Te Ao Mārama 2012
- Digital yearbook collection

Businesses
- Business characteristics
- Business finance
- Business growth and innovation
- Research and development

Economic indicators
- Balance of payments
- Consumers price index (inflation)
- Gross domestic product
- National accounts
- Prices indexes
- Productivity

Education and training
- Primary education
- Secondary education
- Tertiary education
- Schools Corner

Environment
- Natural resources
- Sustainable development

Government finance
- Central government
- Crown research institutes
- District health boards
- Local government

Health
- Abortion
- Disabilities
- Injuries
- Life expectancy

Income and work
- Employment and unemployment
- Income
- Strikes

Industry sectors
- Accommodation
- Agriculture, horticulture, and forestry
- Alcohol availability
- Bioscience and biotechnology
- Construction
- Energy
- Film and television
- Imports and exports
- Information communications and technology
- Manufacturing and production
- Retail trade
- Tourism
- Wholesale trade

People and communities
- Asian peoples
- Children
- Crime and justice
- Families
- Geographic areas
- Households
- Housing
- Language
- Māori
- Marriages, civil unions, and divorces
- Older people
- Pacific peoples
- Time use
- Women

Population
- Births
- Census
- Deaths
- Estimates and projections
- Migration

Questionnaires and Forms

http://www2.stats.govt.nz/domino/external/quest/sddquest.nsf/byName?openview
DDI Data Lifecycle

Concept → Collection → Processing → Distribution → Discovery → Analysis

Archive

Repurposing
Create a single, canonical source for all this information

Solution: central repository
Overview

1. Statistics New Zealand Metadata Architecture
2. Colectica Integration
3. Results
Architecture: Repository

- Central, authoritative source of metadata
- Built on DDI 3, ISO 11179, and Web Service standards
- Full audit trail and provenance tracking
- Item relationships
- Search
- Annotations
- Secure Authentication and Encryption
- Item and Type-based permissions
Colectica Designer is one way to access Colectica Repository.

It’s also easy for programmers to interact with Colectica Repository to build custom tools.

Colectica SDK and Web Services enable this.
Command line tools for batch processing

- MetadataConverter
- SPSSToDDI3
- BlaiseToDDI3
- CASESToDDI3

- DocumentationGenerator
- Validator
- RepositoryImport
- RepositoryExport
Search and browse metadata from Colectica Repository on the Web.

http://midus.colectica.org/
Architecture: Publication Workflow

- Internal Repository
- Workflow Services
- Public Repository

Tools:
- Colectica Designer
- Colectica SDK
- Colectica Toolkit
- Colectica Portal (internal)

Repositories:
- Colectica Repository (Production)
- Colectica Repository (Public)
Architecture: Public Portal

- Colectica Portal (Public)
- Colectica Repository (Public)
- Colectica Workflow Server
  - Colectica Designer
  - Colectica SDK
  - Colectica Toolkit
  - Colectica Portal (internal)
- Colectica Repository (Production)

Available April 2013
Architecture: Not Quite a Complete Picture

Colectica Portal (Public)

Custom Tools

Colectica Repository (Public)

Colectica Workflow Server

Colectica Designer

Colectica SDK

Colectica Toolkit

Colectica Portal (internal)

Colectica Repository (Production)
Architecture: Production and Test Environments

Production

Test
Colectica Integration
Getting to Production

**Option 1**
Build from scratch

**Option 2**

**Goal**
Extend
Extending Colectica

- Versioning and synchronization improvements
- Colectica Portal customization
- Colectica Workflow Services
- Colectica SDK for custom integrations
Integration: Questionnaire Design

**Old:** Questionnaire designers create static flow charts, give to Blaise programmers

**New:** Application team created tool that produces Blaise and pulls questions from Colectica
Integration: Additional Tools

- Data Processing
- Data Dissemination
Results
Key Result 1 – Metadata Capture

“We used to record all metadata at the end of the lifecycle.”

“Now, curators capture the information when they first think of it.”
Key Result 2 - Archiving

- Old Process:
  - Manually mark up DDI 2 XML

- New Process:
  - Information is entered into Colectica
  - A program grabs DDI from Colectica, harvests all information from network drives, ingests into Archive
  - Archivists just have to understand Colectica
  - With content guidelines created by Statistics New Zealand, this is very easy
Key Result 2 - Archiving

Time to Train Archivists

- Old: 3 - 4 Months
- New: 2 Weeks
Facts and Figures

1,008 Datasets
200 Series
20 - 40 Metadata Curators
219 Unique Portal Users
Thank You

- Adam Brown
  adam.brown@stats.govt.nz

- Sally Vermaaten
  sally.vermaaten@stats.govt.nz

- Jeremy Iverson
  jeremy@colectica.com

- Dan Smith
  dan@colectica.com