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

colectica

<ddi>

A Case Study

Authors

Statistics New Zealand

colectica

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

□ Jeremy Iverson
  jeremy@colectica.com

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

□ Dan Smith
  dan@colectica.com

Statistics New Zealand

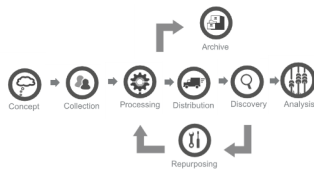❝ "Turning data into relevant knowledge, efficiently."

□ Ensure New Zealand has the statistical information it needs to grown 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

DDI Data Lifecycle

Statistics New Zealand Metadata Infrastructure Project

- Create canonical sources for key information created and used during the statistical business process
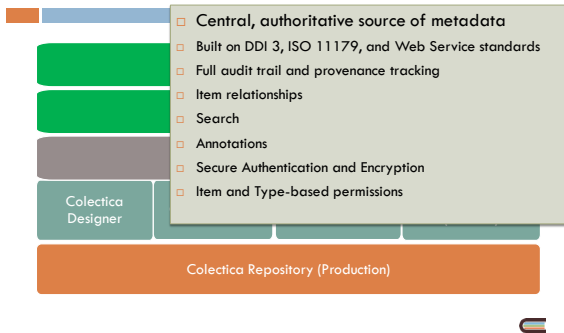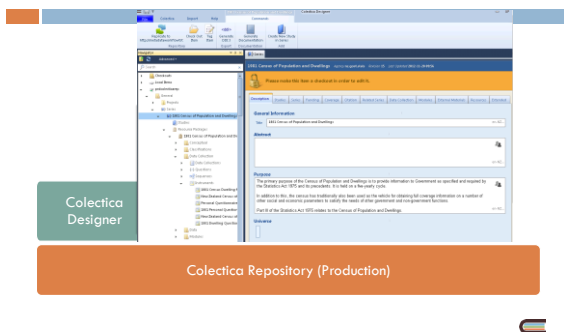- Solution: central repositories

Overview

1. Use of Colectica at Statistics New Zealand
2. Colectica Integration
3. Results

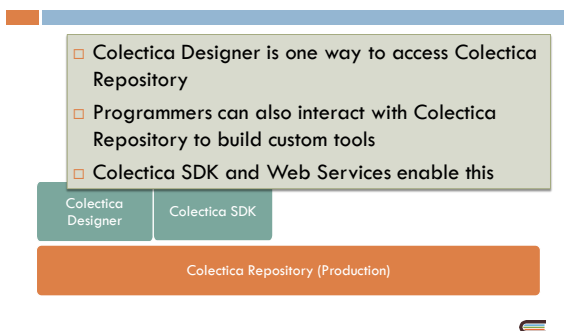## Use of Colectica at Statistics New Zealand

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

Colectica Repository (Production)

Architecture: Designer

Colectica
Designer

Colectica Repository (Production)

Architecture: SDK

- Colectica Designer is one way to access Colectica Repository
- Programmers can also interact with Colectica Repository to build custom tools
- Colectica SDK and Web Services enable this

Colectica
Designer

Colectica SDK

Colectica Repository (Production)

Architecture: Toolkit

☐ Command line tools for batch processing

- **MetadataConverter**
- **SPSSToDDI3**
- **BlaiseToDDI3**
- **CASESToDDI3**

- **DocumentationGenerator**
- **Validator**
- **RepositoryImport**
- **RepositoryExport**
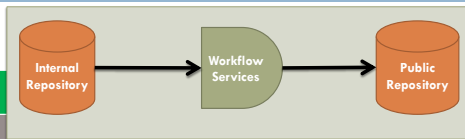
Colectica Designer	Colectica SDK	Colectica Toolkit

Colectica Repository (Production)

Architecture: Portal

Architecture: Publication Workflow

| Internal Repository | → | Workflow Services | → | Public Repository |

Colectica Designer	Colectica SDK	Colectica Toolkit	Colectica Portal (internal)

Colectica Repository (Production)

Architecture: Public Portal

□ Available April 2013	

Colectica Portal (Public)

Colectica Repository (Public)

Colectica Workflow Server

Colectica Designer	Colectica SDK	Colectica Toolkit	Colectica Portal (internal)

Colectica Repository (Production)
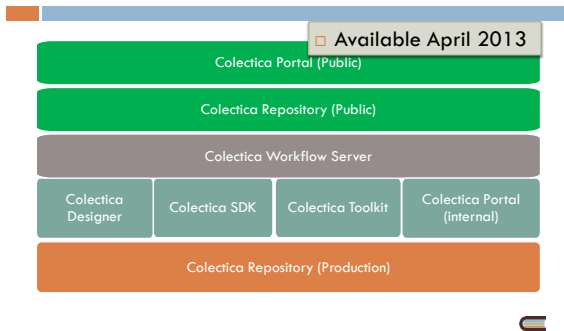
Architecture: Not Quite a Complete Picture

Colectica Portal (Public)

□ Custom Tools	Colectica Repository (Public)

Colectica Workflow Server

Colectica Designer	Colectica SDK	Colectica Toolkit	Colectica Portal (internal)

Colectica Repository (Production)
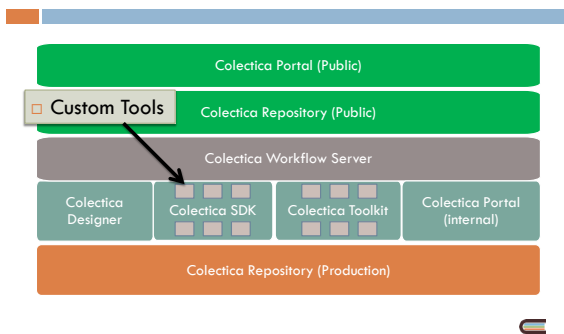
Architecture: Production and Test Environments

Production

Test
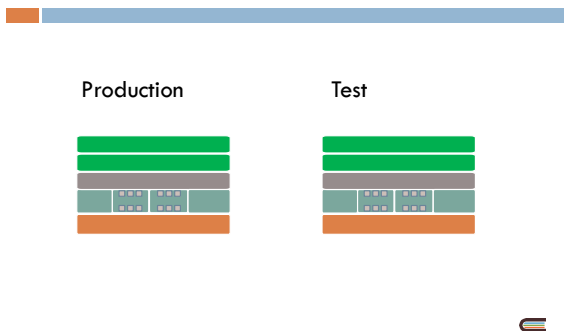
Colectica Integration

Getting to Production

Option 1   **Goal**

Build from scratch

**Option 2**   **Goal**

Extend

Extending Colectica

- ☐ Versioning and synchronization improvements
- ☐ Item-level and item-type permissions
- ☐ Repository usage statistics
- ☐ Colectica Portal customization
- ☐ Colectica Workflow Services

- ☐ Colectica SDK for custom integrations

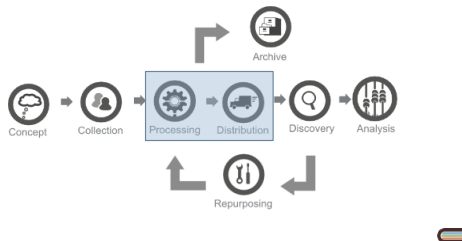## Integration: Questionnaire Design

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

- **New**: Application team created a prototype tool that pulls questions from Colectica, allows design of question flow, and creates Blaise
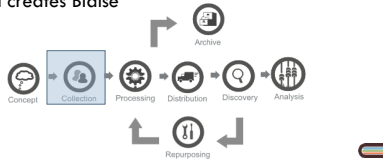
Integration: Additional Tools

- Data Processing
- Data Dissemination

Results

Key Result 1 – Metadata Capture

> "We used to record all metadata at the end of the lifecycle."

> "Now, survey designers 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
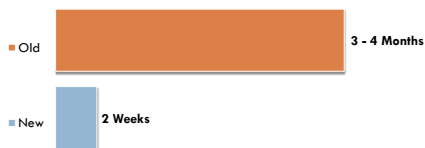
Facts and Figures

1,008
Datasets

200
Series

20 - 40
Metadata
Creators

219
Unique
Portal Users

Future of Colectica and NSOs

□ Improving support for GSIM

□ Continue to improve usability of the Colectica
interface for those unfamiliar with DDI

□ Continue to work with statistical agencies to tailor
Colectica to their needs and meet their goals of
modernisation

Thank You

Statistics
New Zealand
TATAURANGA AOTEAROA

colectica

□ Adam Brown
adam.brown@stats.govt.nz

□ Jeremy Iverson
jeremy@colectica.com

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

□ Dan Smith
dan@colectica.com