

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



A Case Study

# Authors

2



□ Adam Brown

adam.brown@stats.govt.nz

□ Sally Vermaaten

sally.vermaaten@stats.govt.nz



□ Jeremy Iverson

jeremy@colectica.com

□ Dan Smith

dan@colectica.com



# Statistics New Zealand



**“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

## Name of Questionnaire

[Accommodation Survey 2004 \(AS/LF\(T\), AS/LF\) - Accommodation 2004, long form](#)

[Accommodation Survey 2004, short form \(AS/SF\(T\), AS/SF\(C\)\) - Accommodation 2004, short form](#)

[Accommodation Survey 2005, long form \(AS/LF\(T\), AS/LF\) - Accommodation 2005, long form](#)

[Accommodation Survey 2005, short form \(AS/SF, AS/SF\(T\)\) - Accommodation Survey, short form](#)

[Accommodation Survey 2006 \(AO/SQ/01\) - Scoping Questionnaire](#)

[Accommodation Survey 2007 \(AO/BI/01\) - Accommodation](#)

[Accommodation Survey 2010 \(AO/BI/01\) - Accommodation](#)

[Accommodation Survey \(date unknown\) \(AO/SF/01\) - Shortform](#)

[Accommodation Survey \(date unknown\) \(AO/SQ/01\) - Scoping Questionnaire](#)

[Agricultural Production Census 2002 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Census 2007 \(AP/EP/01\) - agriculture, census, livestock, horticulture, forestry, farming](#)

[Agricultural Production Census 2012 \(AP/EP/01\) - Agriculture, Livestock, Arable, Forestry, Horticulture](#)

[Agricultural Production Survey 1996 \(AP/FS/LH\) - Livestock, Horticulture](#)

[Agricultural Production Survey 1999 \(AP/EB/06\) -](#)

[Agricultural Production Survey 2000 \(AP/EB/07\) - Horticulture](#)

[Agricultural Production Survey 2003 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Survey 2004 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Survey 2005 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Survey 2006 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Survey 2008 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Survey 2009 \(AP/EP/01\) - Agriculture, Horticulture, Forestry](#)

[Agricultural Production Survey 2010 \(AP/EP/01\) - Agriculture, Livestock, Arable, Forestry](#)

[Agricultural Production Survey 2011 \(AP/EP/01\) - Agriculture, Arable, Livestock, Forestry, Horticulture](#)

[Annual Enterprise Survey 2007/2008 \(AF/AC/01\) - Services to Agriculture](#)

[Annual Enterprise Survey 2007/2008 \(AF/AC/PF\) - Services to Agriculture](#)

[Annual Enterprise Survey 2007/2008 \(AF/AL/1E\) - Summary of Accounting Divisions](#)

[Annual Enterprise Survey 2007/2008 \(AF/AL/B\) - Statement of Financial Position](#)

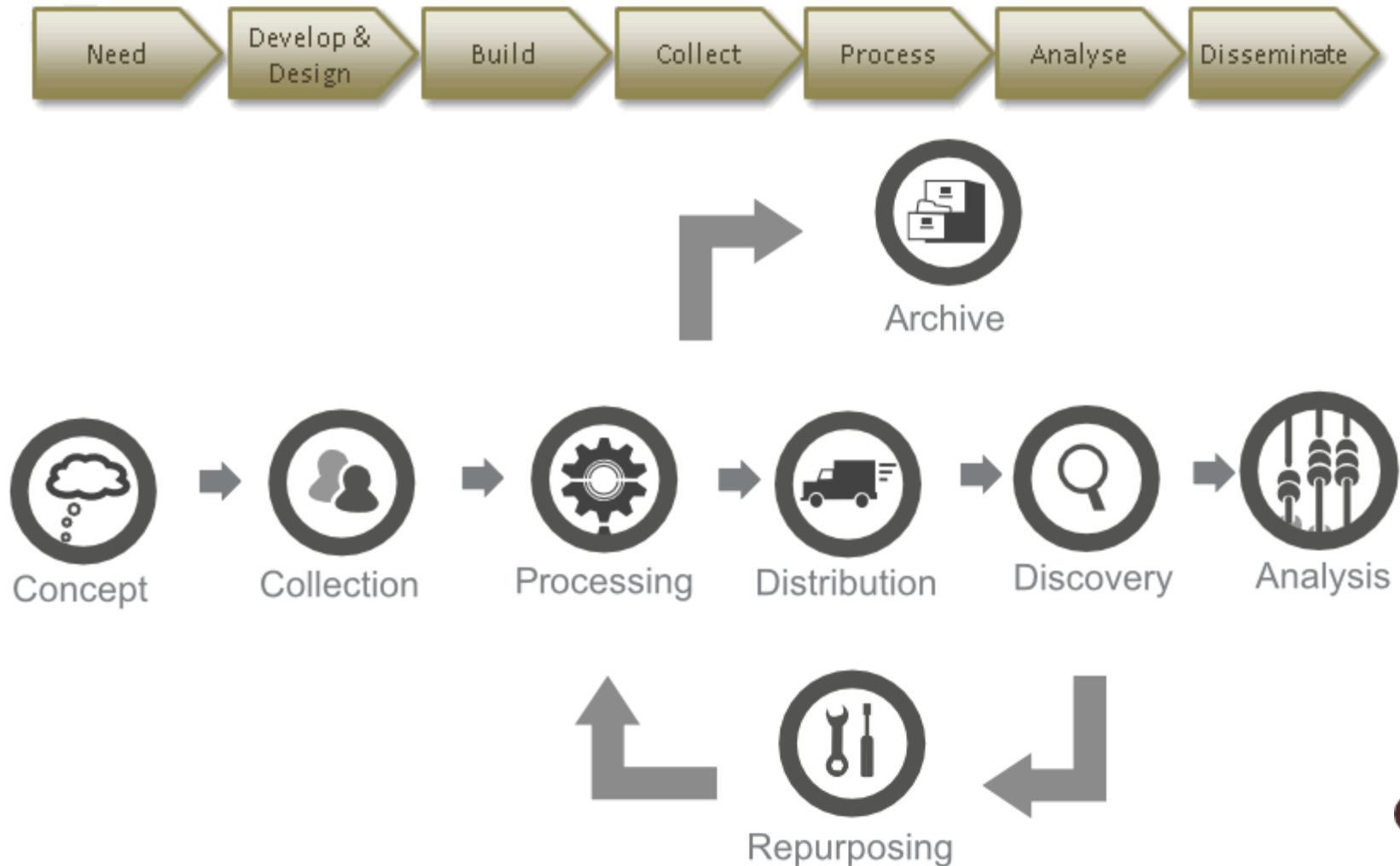
[Annual Enterprise Survey 2007/2008 \(AF/BC/01\) - Construction](#)

[Annual Enterprise Survey 2007/2008 \(AF/BC/PF\) - Construction](#)

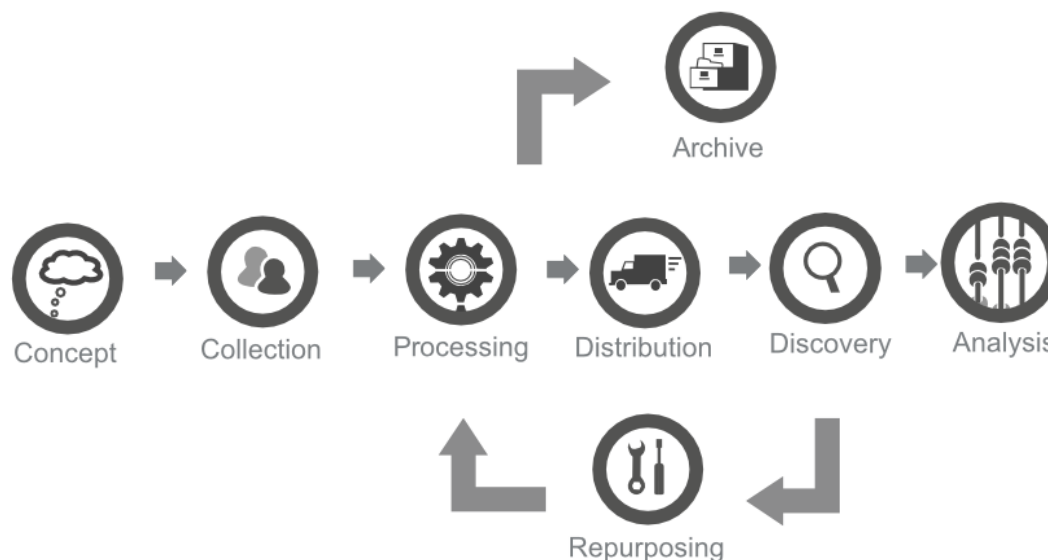
[Annual Enterprise Survey 2007/2008 \(AF/BT/01\) - Building Trades](#)

# DDI Data Lifecycle

6



# 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

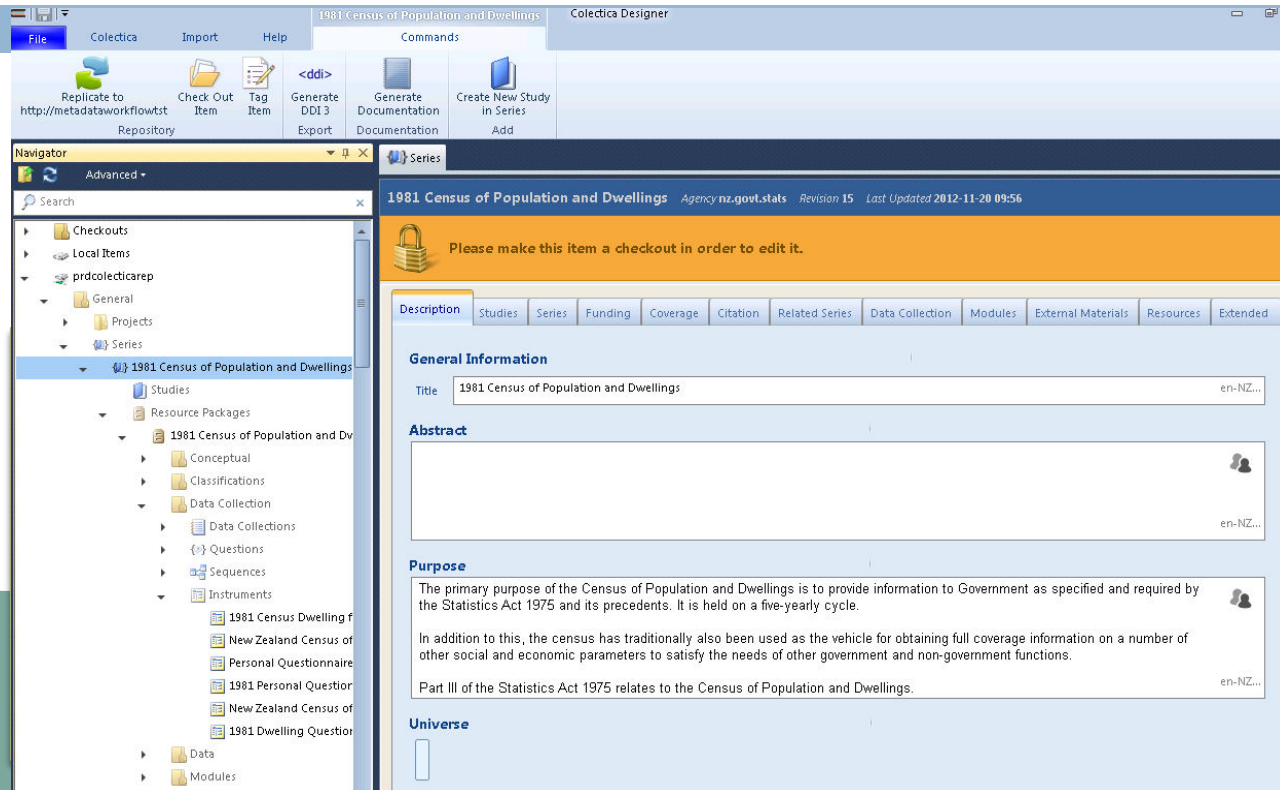
- Central, authoritative source of metadata
- Built on DDI 3, ISO 11179, and Web Service standards
- Full audit trail and provenance tracking
- Item relationships
- Search
- Annotations
- Secure Authentication and Encryption
- Item and Type-based permissions

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

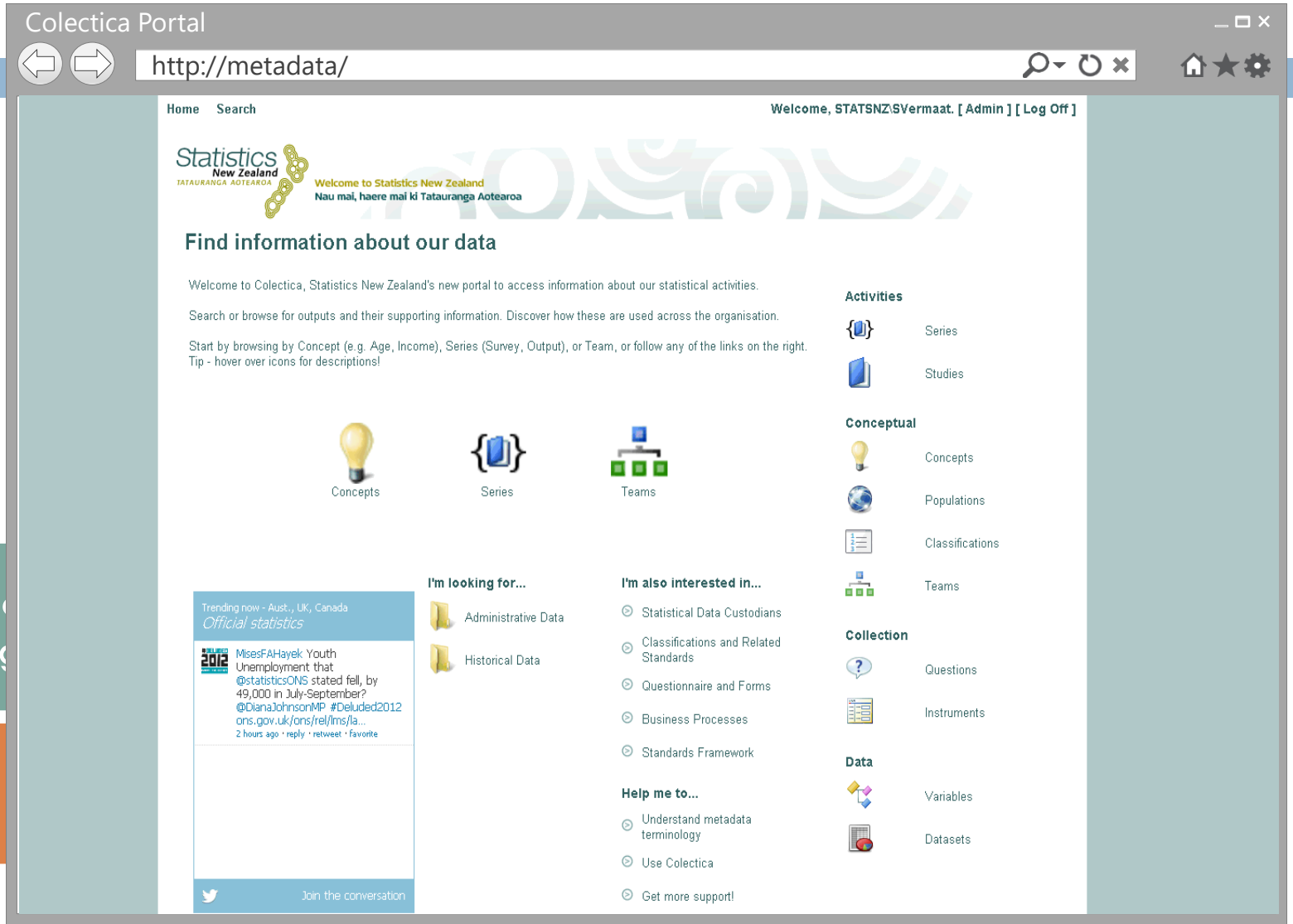
Colectica SDK

Colectica Toolkit

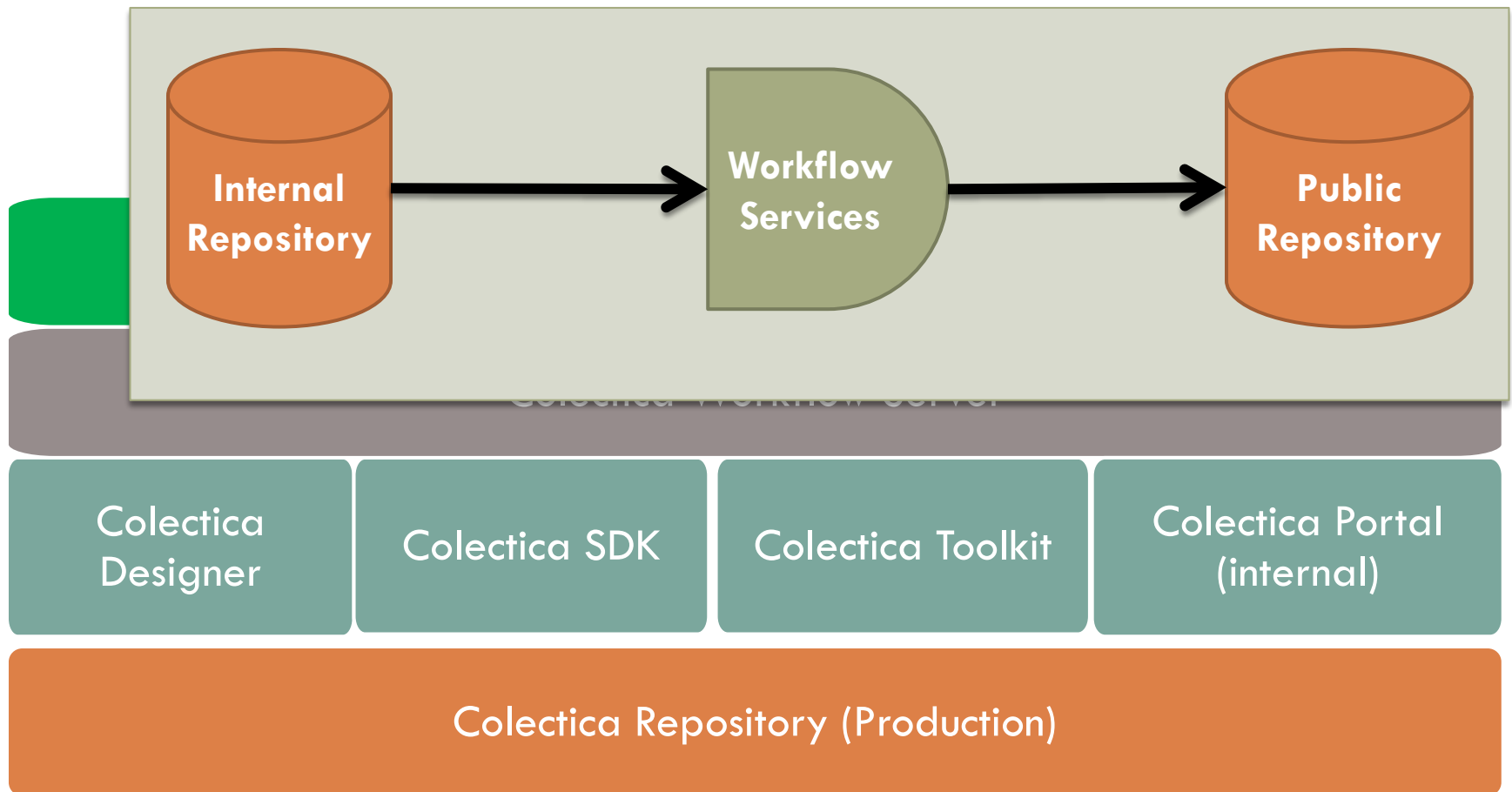
Colectica Repository (Production)



# Architecture: Portal



# Architecture: Publication Workflow



# Architecture: Public Portal

□ Available April 2013

Colectica Portal (Public)

Colectica Repository (Public)

Colectica Workflow Server

Colectica  
Designer

Colectica SDK

Colectica Toolkit

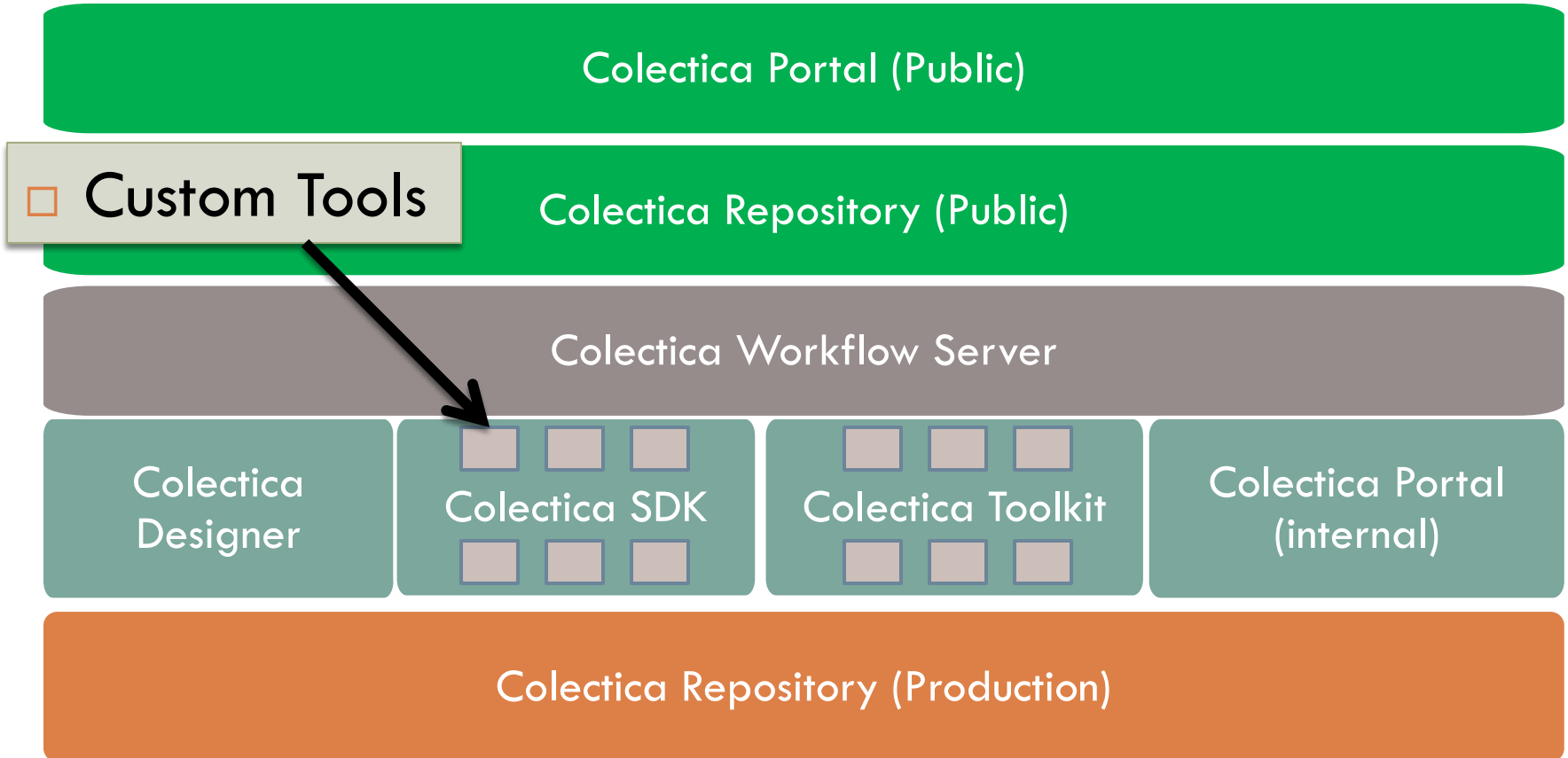
Colectica Portal  
(internal)

Colectica Repository (Production)





# Architecture: Not Quite a Complete Picture

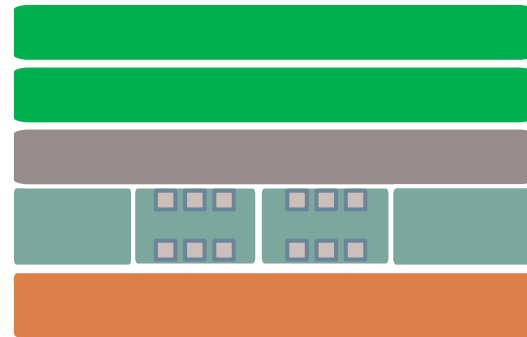


# Architecture: Production and Test Environments

## Production



## Test





# Colectica Integration

# Getting to Production

## Option 1

**Goal**



Build from scratch

## Option 2

**Goal**



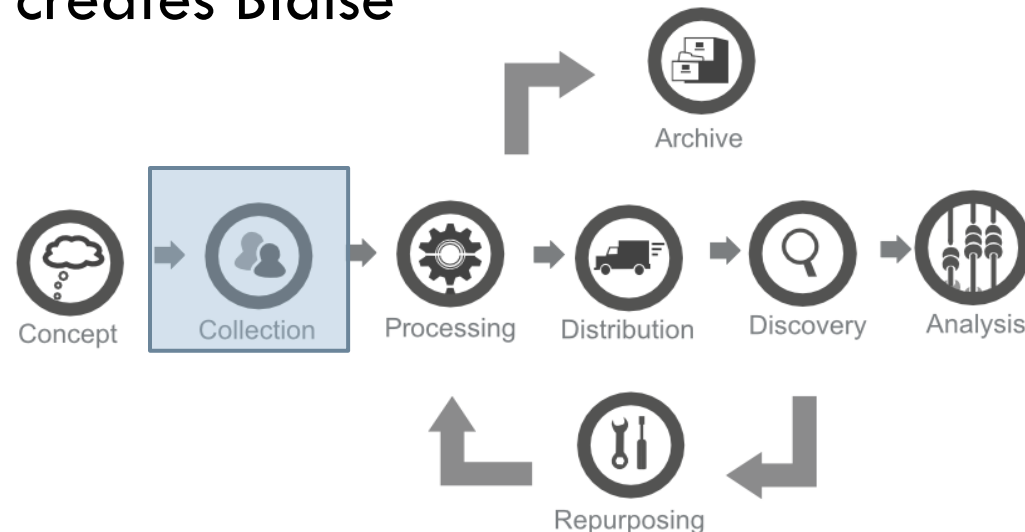
# Extending Colectica

- ❑ Versioning and synchronization improvements
- ❑ Item-level and item-type permissions
- ❑ Repository usage statistics
- ❑ Colectica Portal customization
- ❑ Colectica Workflow Services
  
- ❑ Colectica SDK for custom integrations



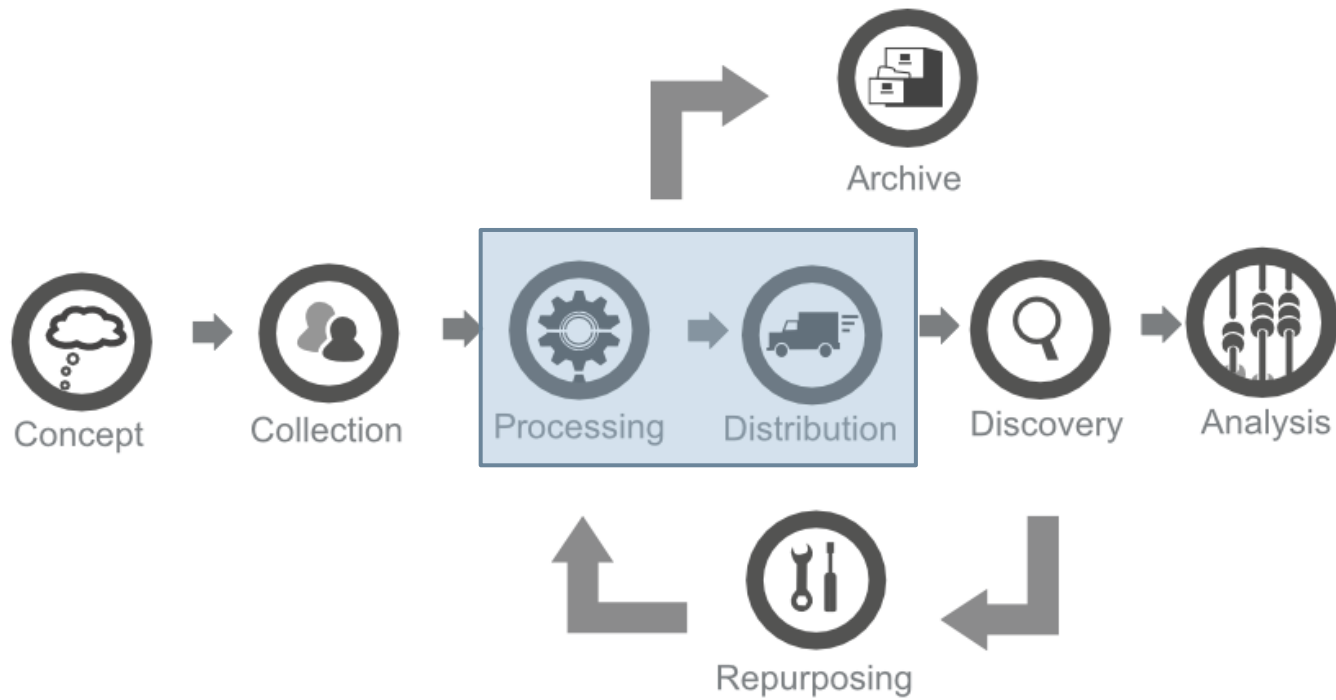
# Integration: Questionnaire Design

- **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

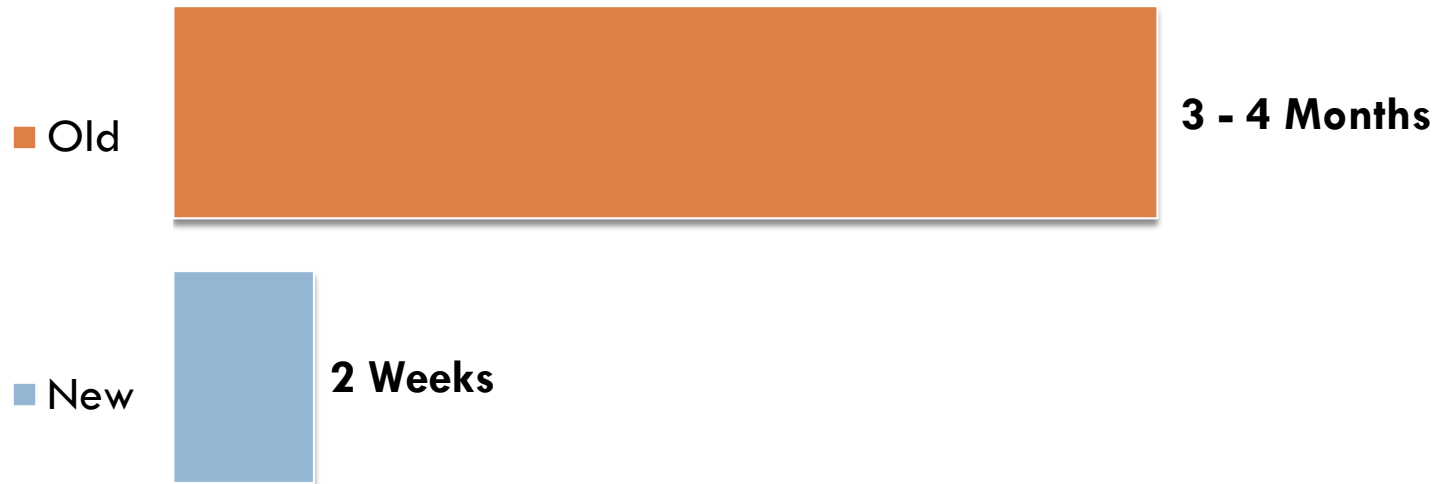
## □ New Process:

- ▣ 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



# 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

30



□ Adam Brown

adam.brown@stats.govt.nz

□ Sally Vermaaten

sally.vermaaten@stats.govt.nz



□ Jeremy Iverson

jeremy@colectica.com

□ Dan Smith

dan@colectica.com

