Integrated Data Infrastructure (IDI)
Unleashing the power of data through integration

Sarah Minson
Operations Strategy & Development Senior Manager

October 2016

Unleashing the power of data to change lives
Integrated Data Infrastructure (IDI)

An integrated database containing confidentialised longitudinal microdata about people, households, and businesses

- Student loans & allowances
- Education
- Benefits
- Migration & movements
- Person-centred data
- Tax
- Families & households
- Justice
- Health & safety
- Business data
Data in the IDI

**Benefits**
- Benefit data – from 1990
- Child benefit, and 2nd and 3rd tier benefit payments – From 1993

**Education**
- Early Childhood Education Providers
- Primary Education – from 2007
- Secondary Education – From 2004
- Tertiary Education – From 1994

**Justice**
- Charges – From 1992
- Sentencing – From 1998
- Victims – From 1998
- Offenders – From 2014

**Census**
- 2013 Census

---

**Health and safety**
- Popular Cohort Demographics – From 2004
- Chronic conditions – From 1945
- Cancer registrations – From 1995
- Mortality
- Pharmaceuticals – From 2005
- Publicly Funded Hospital Discharges – from 1988
- National non-admitted patient collection – From 2007
- Before school checks – 2011-2013
- PRIMHD – 2008-2013

---

Unleashing the power of data to change lives
Spine created by linking births, tax, and visa data together

Other sources (aka nodes) are linked one-to-one with the spine.

Unleashing the power of data to change lives
Five safes

- People
- Projects
- Settings
- Data
- Output
Access to IDI for research

**Go**

- Research is for a statistical purpose
- Research is for the public good
- Research will be conducted by a credible team
- Suitable data is available
- Statistics NZ can enforce an agreement

**Check**

- Research is for commercial gain
- Research is for regulatory purposes
- Research is for service provider evaluation
- Quality of data is not adequate for research purpose
- Lead researcher is a student or junior staff member
- Researchers are based overseas
- Unrecognised research institution
- Data integration is required

**Stop**

- Research is about named people or businesses
- Researchers do not have organisational support
- Research will not be released
- Complex legislative issues affect access

You need to resolve all red issues before microdata access can be considered.

You need five greens to consider getting access to microdata for research.

You need to talk to us to check each orange issue for potential problems.

Unleashing the power of data to change lives
Crime

How can crime and harm from crime be prevented?
Outcomes of tertiary study

What should I study?

Temporary migration

Should international students be allowed longer work entitlements?
Teenage Parents

Do Teen Parent Units result in increased school qualifications for teenage mothers?
## What do the public think?

Moving from naïve trust to informed trust

<table>
<thead>
<tr>
<th>Less acceptable</th>
<th>More acceptable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambiguous and general</td>
<td>Purpose</td>
</tr>
<tr>
<td>Insecure</td>
<td>Security</td>
</tr>
<tr>
<td>Open access</td>
<td>Access</td>
</tr>
<tr>
<td>Public/personalised</td>
<td>Privacy</td>
</tr>
<tr>
<td>Private</td>
<td>Interests</td>
</tr>
<tr>
<td>Subjective and biased</td>
<td>Data &amp; Metrics</td>
</tr>
<tr>
<td>Unfair and harmful</td>
<td>Outcomes</td>
</tr>
</tbody>
</table>

- **Purpose**: Transparent and specific need
- **Security**: Secure
- **Access**: Strict restrictions, procedures, and protocols
- **Privacy**: Confidential/anonymous
- **Interests**: Public
- **Data & Metrics**: Representative and meaningful
- **Outcomes**: Fair and beneficial

Unleashing the power of data to change lives
Unleashing the power of data to change lives

**Steward**
Steward data now and for the future

**Provider**
Provide independent and trusted data

**Enabler**
Enable New Zealand decision makers

**Innovator**
Innovate to drive value for customers
Questions