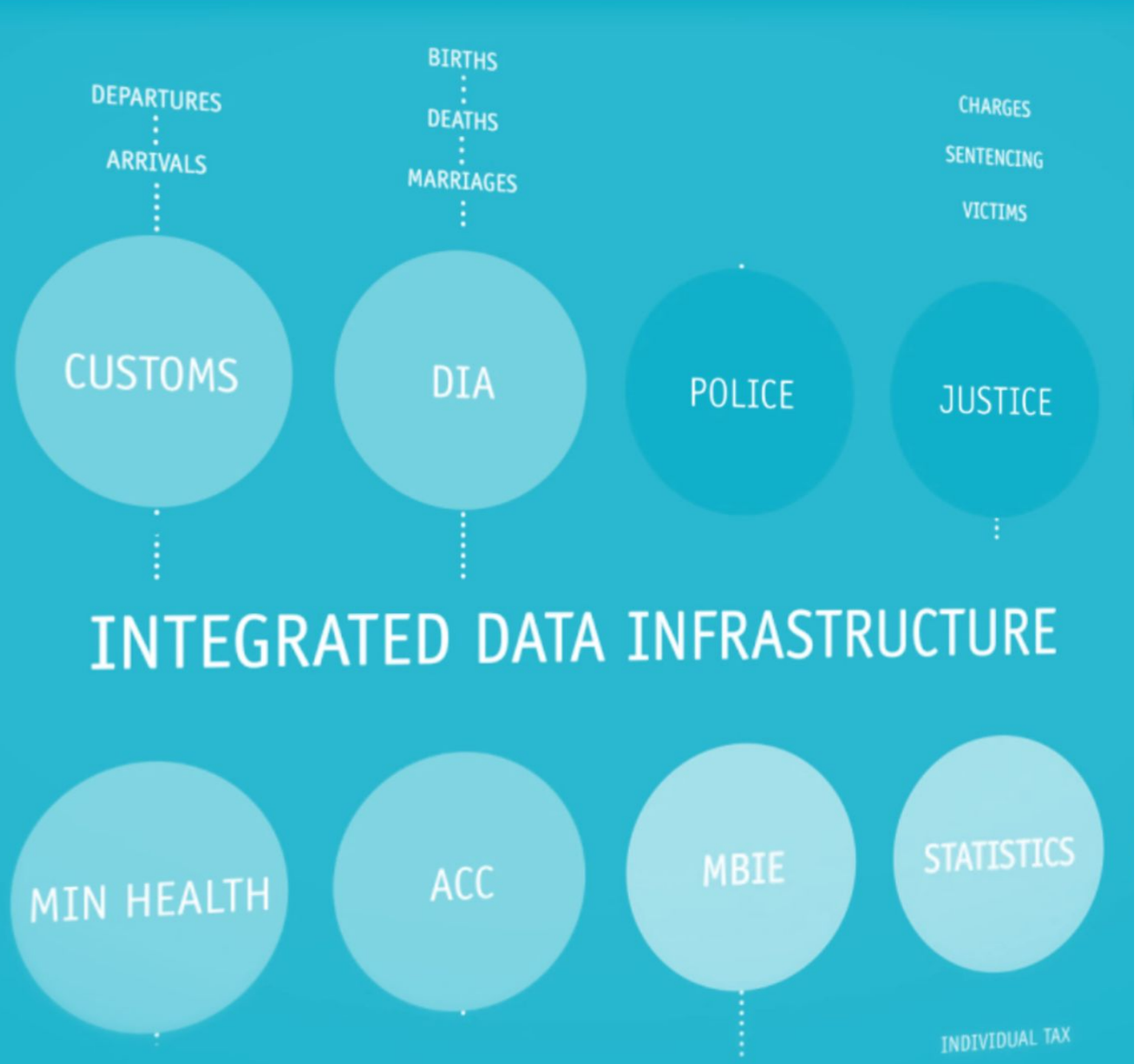
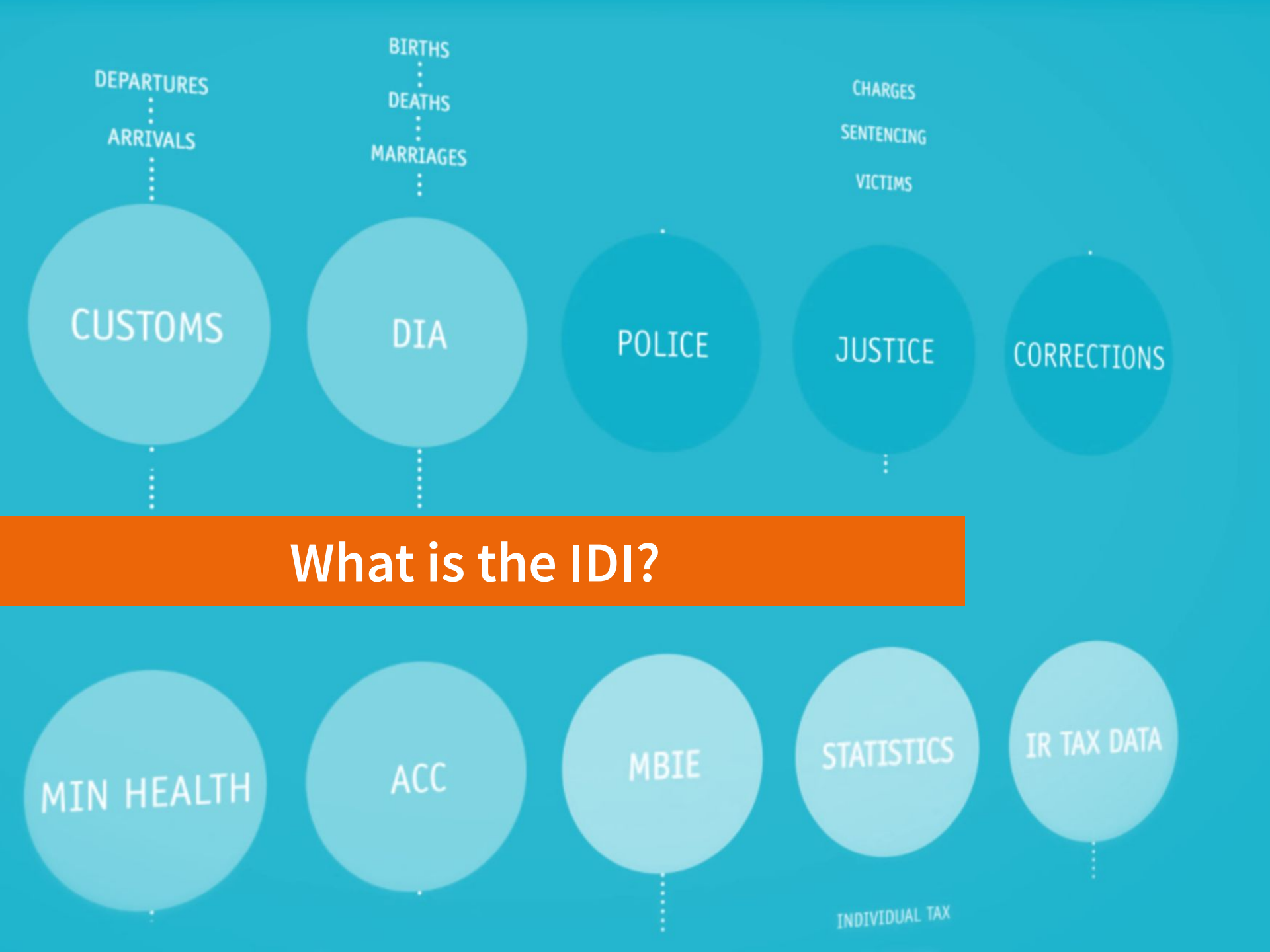


Stats NZ
Tatauranga Aotearoa

Introduction to the IDI

- What is the IDI?
- How is the IDI being used?
- How does it work?
- Technical details
- Privacy by design
- Where to next?

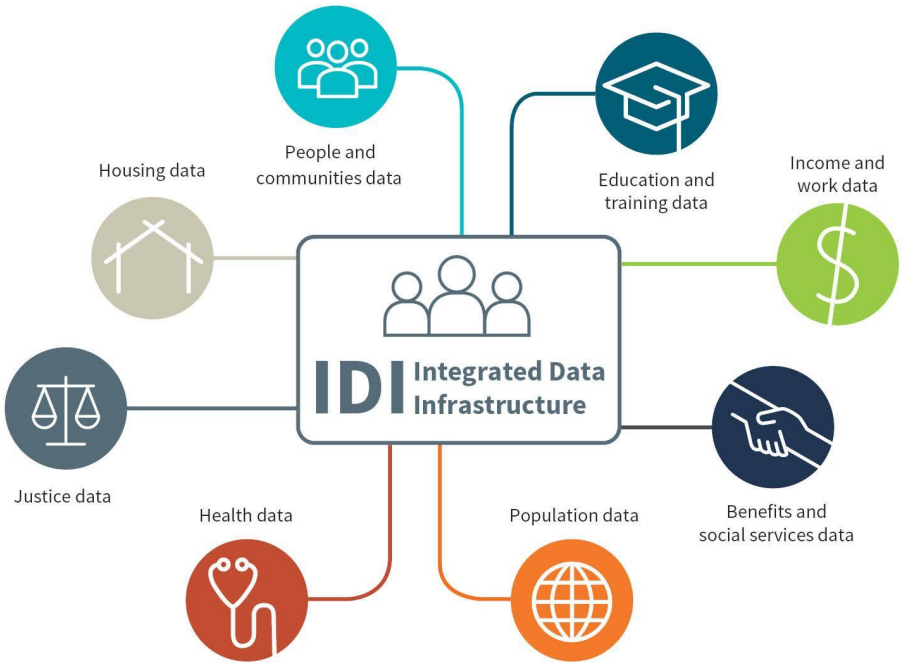

DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
CHARGES
SENTENCING
VICTIMS
POLICE
JUSTICE
CORRECTIONS
What is the IDI?
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

History of the IDI

- 1997 Cabinet endorsed Stats NZ cross agency integrations
- 2011 IDI prototype
- 2012 – Focus on Outcomes
- 2013 Cabinet agree to expand IDI
- 2016 additional funding to improve infrastructure

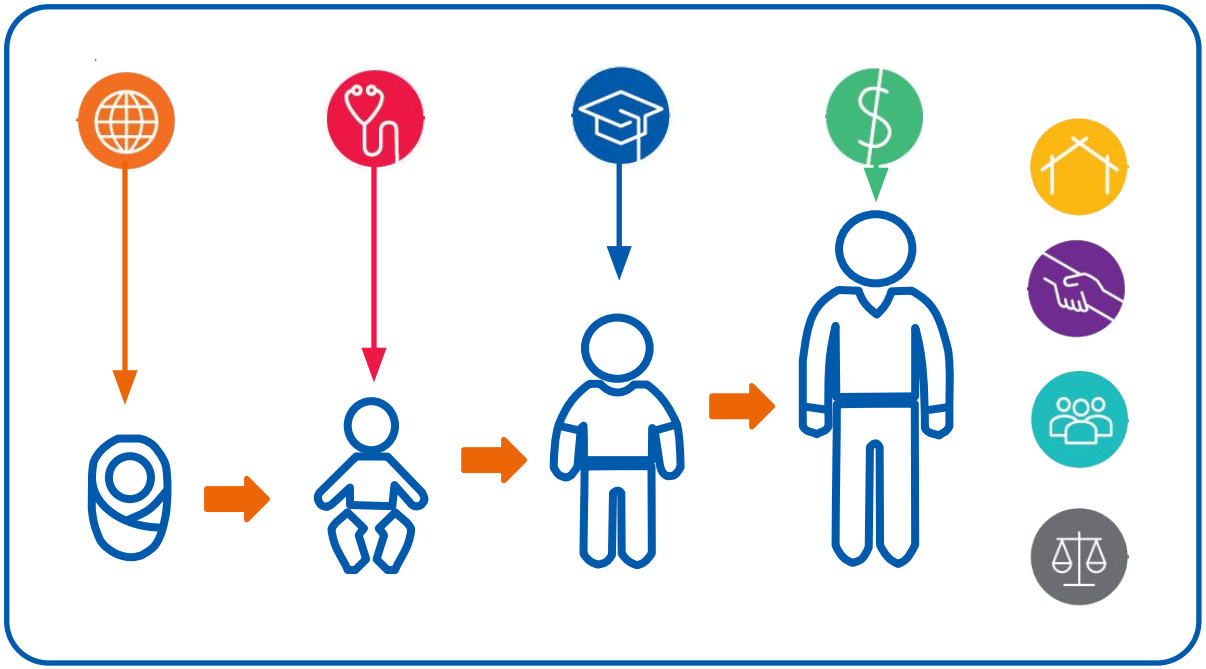
Integrated Data Infrastructure (IDI)

An integrated database containing de-identified longitudinal microdata about people, households, & businesses.



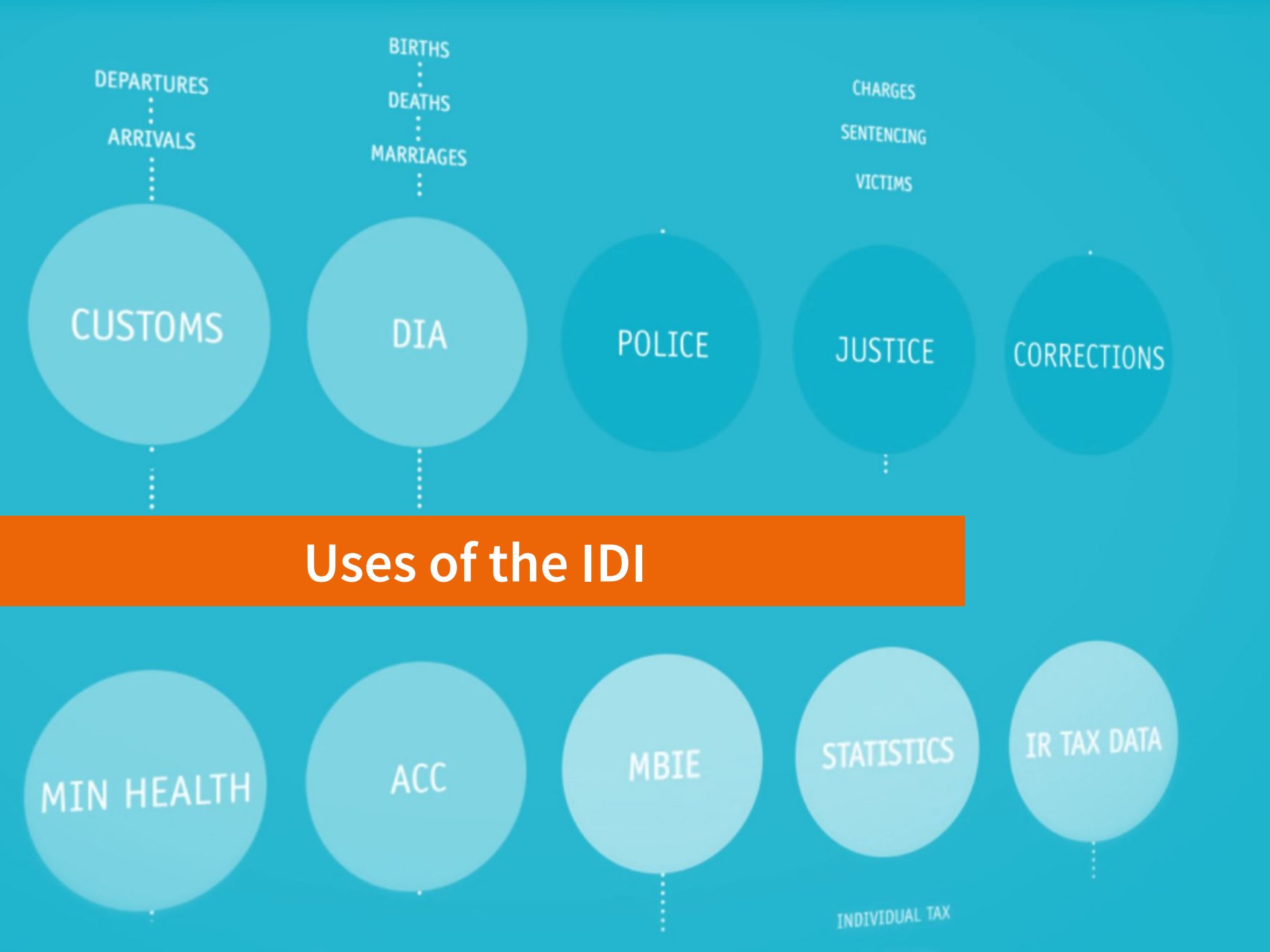
For a full list of data, go to www.stats.govt.nz/idi

The IDI builds a picture of people's lives …



A World Leader

"Internationally, only a handful of administrative datasets are on par with the IDI."


DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
CHARGES
SENTENCING
VICTIMS
POLICE
JUSTICE
CORRECTIONS
Uses of the IDI
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

Angela and Jaxon

How the IDI helps shine a light on…children's health

Are kids in green neighbourhoods less likely to develop asthma?

Parents use of antibiotics and kids obesity

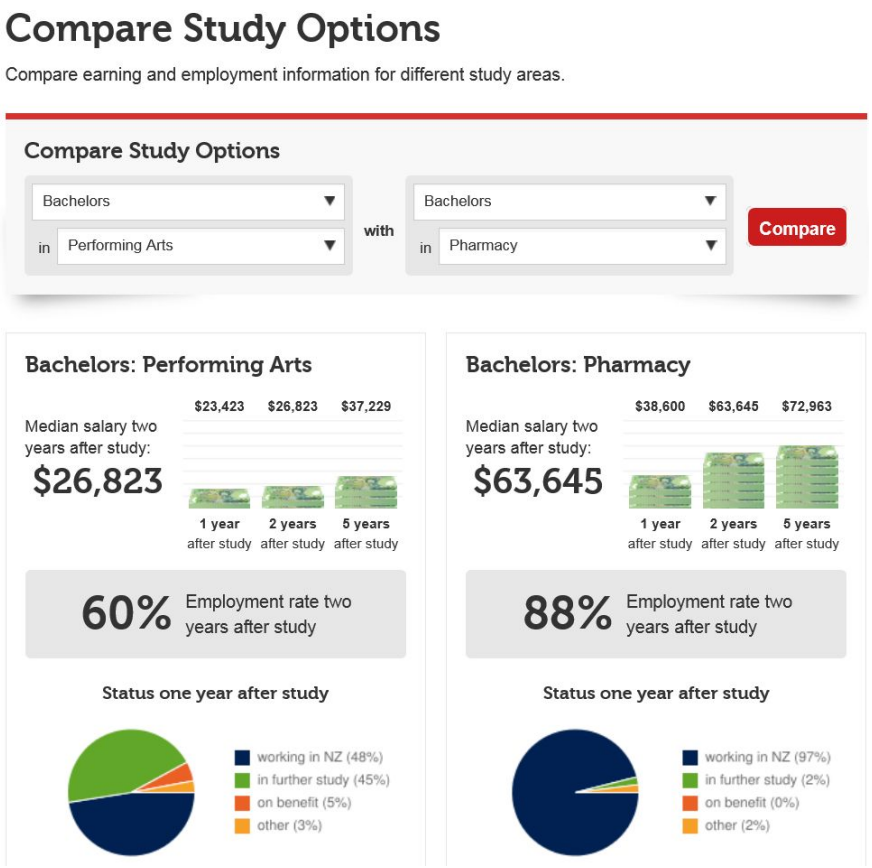
What are the factors for Sudden Unexplained Death in Infancy?

Does a free breakfast help children's dental health?

Helping parents know outcomes for premature babies

Outcomes of tertiary study

What should I study?



Investment Approach to Justice

How can crime and harm from crime be prevented?

Image source: Mediaworks

Cost of Cardiovascular Disease

What is the cost of cardiovascular disease on our health system?

NGOs and iwi

What can Statistics NZ do as kaitiaki tatauranga to better partner with iwi?

Auckland City Mission

How do we improve outcomes for the homeless?

Temporary migration

Should international students be allowed longer work entitlements?

Exploring the gender pay gap

"We used multiple robust methods to find similar outcomes, which was strong evidence for policy makers to act."

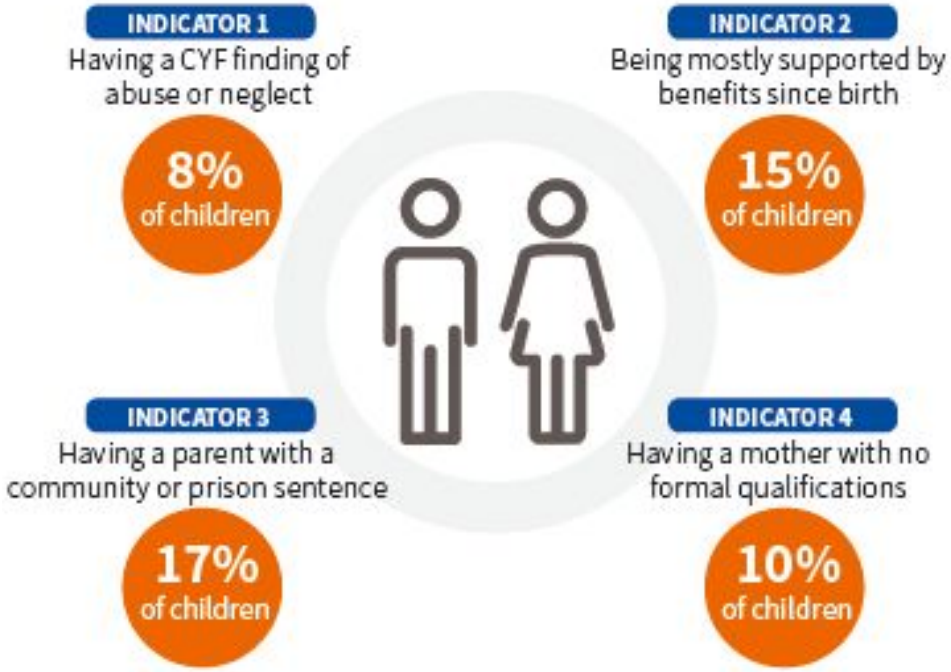
Does becoming a mother impact wages for women?

http://women.govt.nz/documents/closing-gender-pay-gap-actions-employers

Mental Health Insights

What is the impact of mental health on education, employment and training?

Childhood risk factors

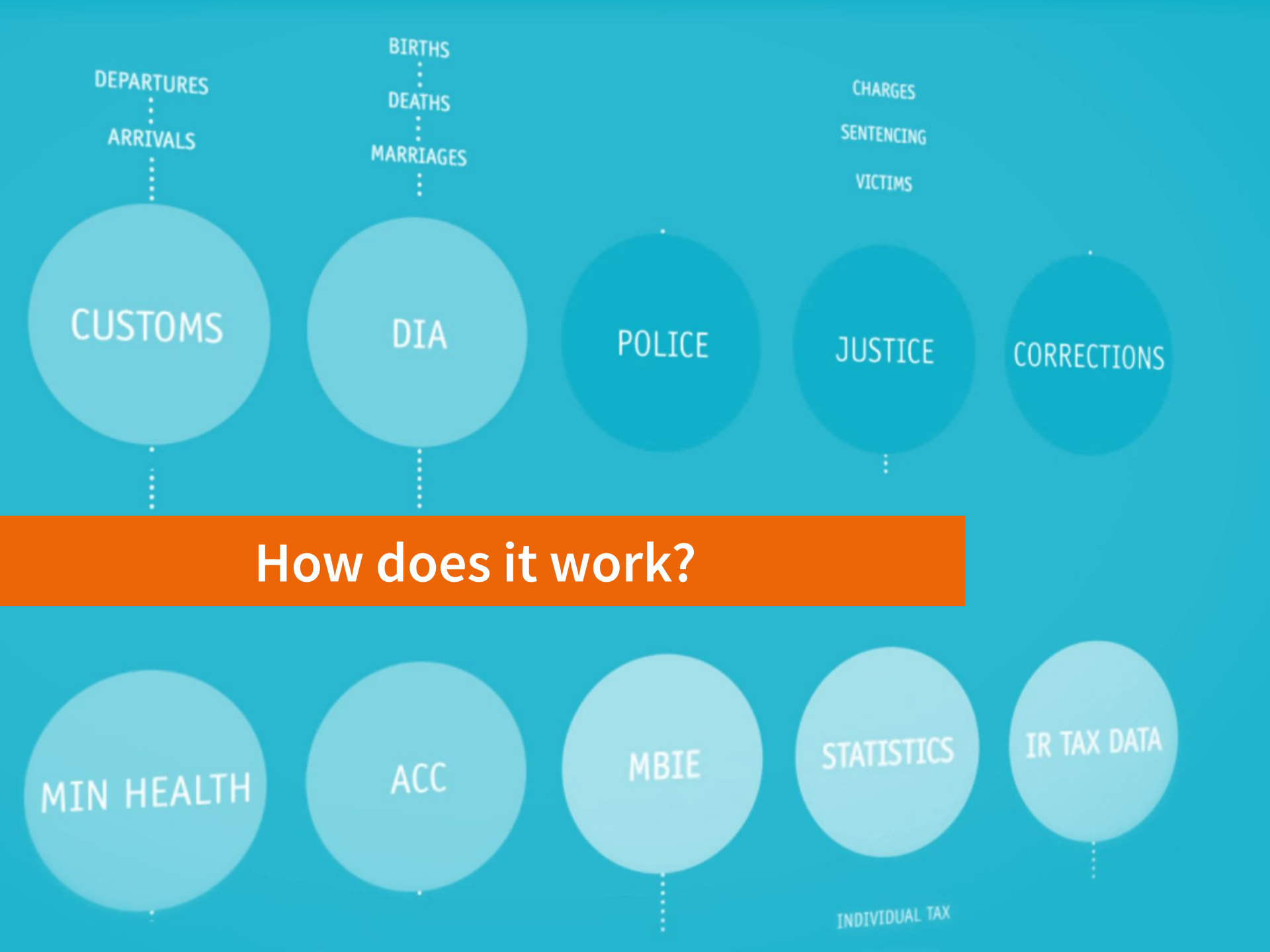


How can we best help children in need?

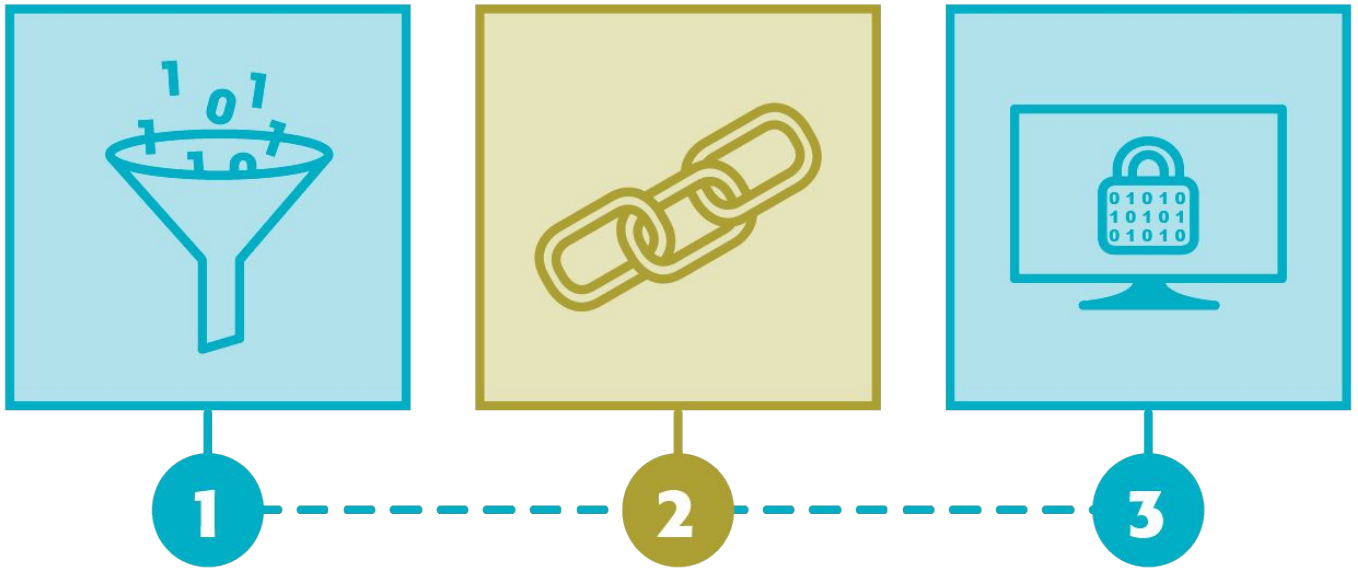
https://insights.apps.treasury.govt.nz/

NEETs

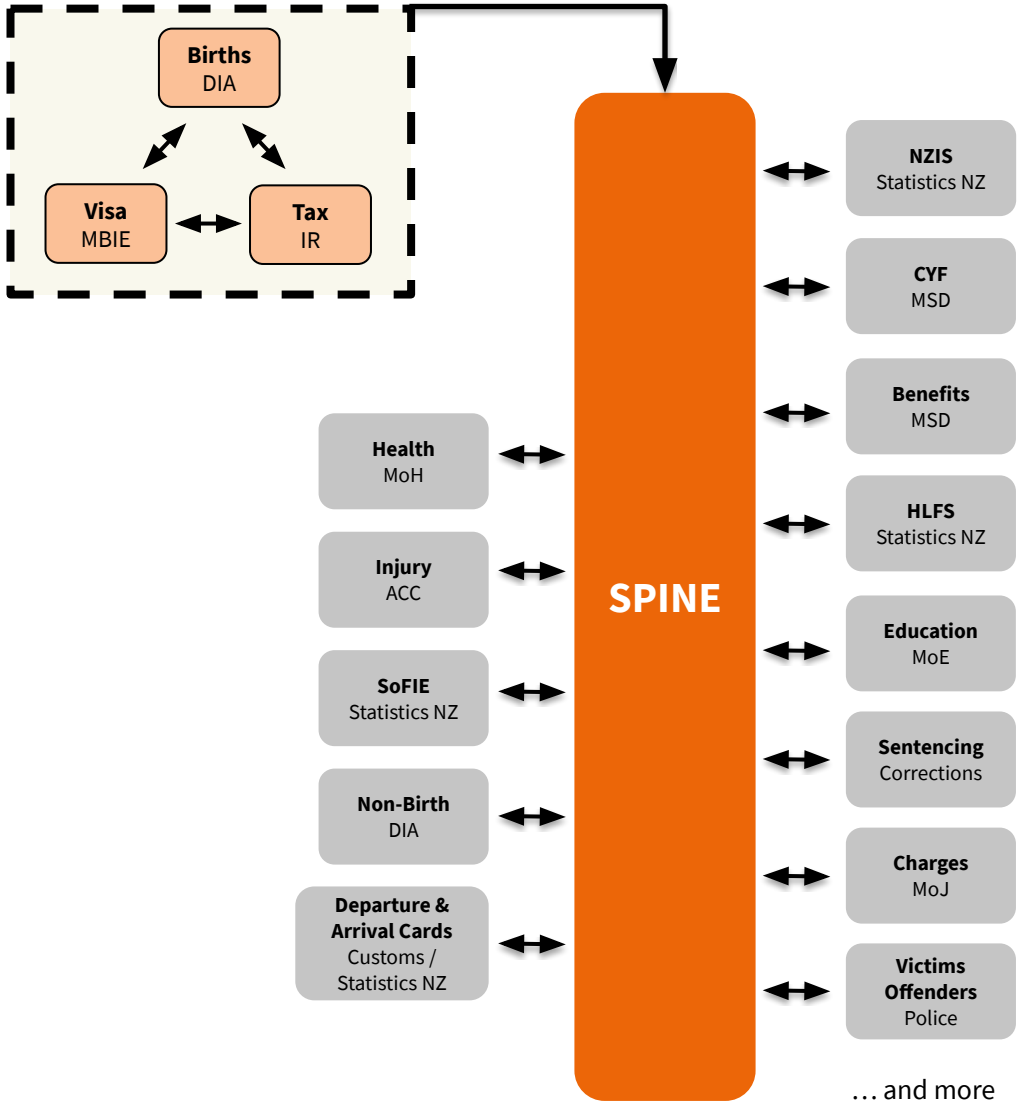
What happens to school leavers who don't go on to further education or employment?


DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
CHARGES
SENTENCING
VICTIMS
POLICE
JUSTICE
CORRECTIONS
How does it work?
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

Flow of data in the IDI



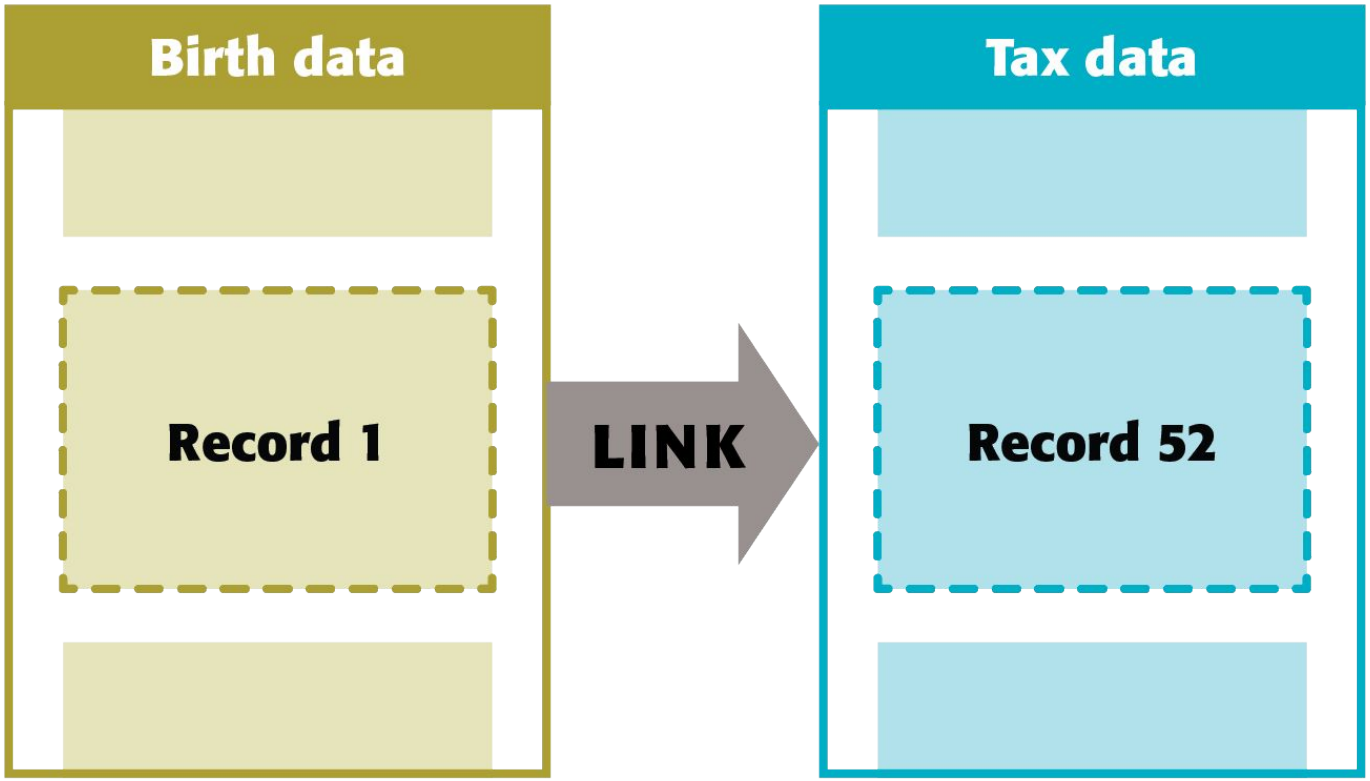
Linking datasets together



1. Spine created by linking births, tax, and visa data together

2. Other sources (aka nodes) are linked with the spine.

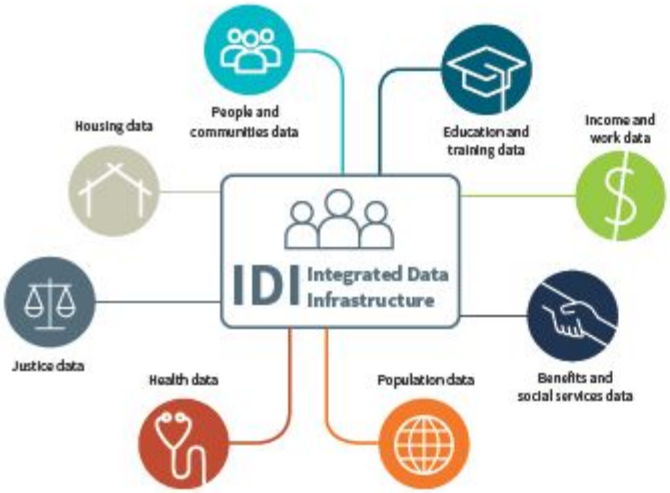
Linking records



Data in the IDI June 2018

Stats NZ Tatauranga Aotearoa

Stats NZ's Integrated Data Infrastructure (IDI) is a large research database containing de-identified microdata about people and households.



The IDI contains person-centred microdata from a range of government agencies, Stats NZ surveys including the 2013 Census, and non-government organisations. For more information about data in the IDI, see **www.stats.govt.nz/integrated-data/integrated-data-infrastructure**

The Longitudinal Business Database (LBD) complements the IDI with microdata about businesses. For more information about data in the LBD, see **www.stats.govt.nz/integrated-data/longitudinal-business-database**

Health data

- B4 School Checks – from 2011
- Cancer registrations – from 1995
- Chronic conditions – from 2007
- General medical services claims – from 2002
- Health tracker – 2006-13
- Laboratory claims – from 2003
- Mortality – from 1988
- Immunisation – from 2006
- National non-admitted patient collection – from 2007
- Pharmaceuticals – from 2005
- PHO enrolments – from 2003
- Population cohort demographics and addresses – from 2004
- Mental health and addiction – from 2008
- Publicly funded hospital discharges – from 1988
- National Needs Assessment and Service Coordination Information System (SOCRATES)
- Maternity – from 2003

Education and training data

- Early childhood education participation – from 2008
- Primary education – from 2007
- Secondary education – from 2004
- Tertiary education – from 1994
- Industry training – from 2001
- Targeted training – from 2001
- Adult competency assessments – from 2014

Benefits and social services data

- Benefits – from 1990
- Youth services – from 2004
- Children's Action Plan – from 1996
- Working for Families – from 2003
- Child, Youth, and Family – from 1991
- Student loans and allowances – from 1992
- ACC injury claims – from 1994
- Family Start – from 2008

Justice data

- Recorded crime: offenders – from 2009
- Recorded crime: victims – from 2014
- Court charges – from 1992
- Sentencing and remand – from 1998

People and communities data

- Auckland City Mission – from 1996
- Migrant Survey – from 2012
- Driver licence and motor vehicle registers
- Longitudinal Immigration Survey of NZ – 2005-09
- General Social Survey – 2008-2016
- Disability Survey – 2013
- Te Kupenga – 2013

Population data

- Border movements – from 1997
- Visa applications – from 1997
- Departure and arrival cards – from 1997
- 2013 Census
- Births, deaths, marriages, and civil unions – from 1840

Income and work data

- Tax and income – from 1999
- NZ Income Survey – from 2006
- Household Labour Force Survey – from 2006
- Survey of Family, Income, and Employment – 2002-10
- Household Economic Survey – from 2006

Housing data

- Tenancy – from 2000
- Social housing – from 1980

Stats NZ operates a five-safes environment, balancing privacy and confidentiality with data insights.

For information about applying to use the IDI or to learn about how we keep the data safe, see **www.stats.govt.nz/integrated-data**

DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
CHARGES
SENTENCING
VICTIMS
POLICE
JUSTICE
CORRECTIONS
Technical details
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

IDI Spine

- Spine:
  - a primary dataset that other datasets are linked to
  - complete list of uniquely identified members of the target population
  - maximum coverage from a minimum number of sources
  - minimum duplicates
  - high quality (low false positive rate and low false negative rate)

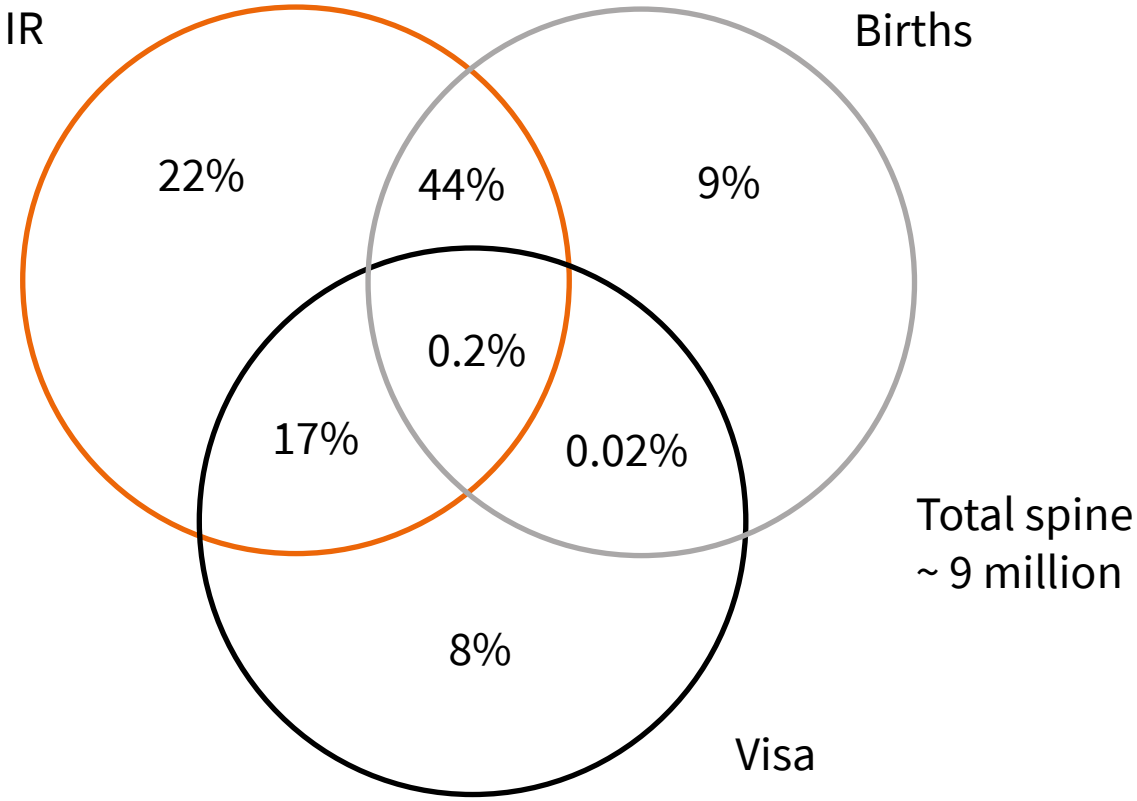
Prototype spine

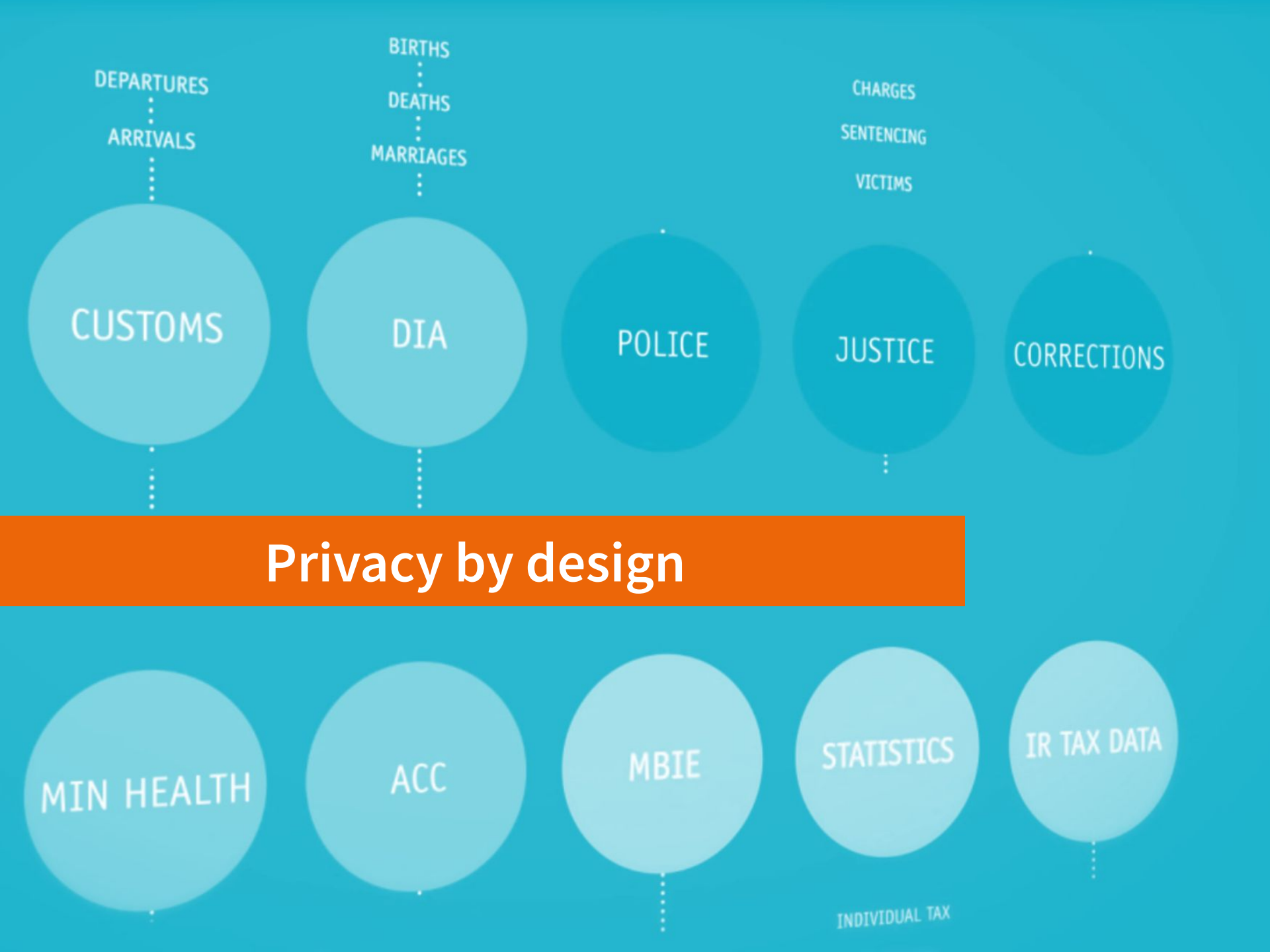
- The target population for the prototype spine is broadly an ‘ever-resident’ population
- It includes
    - people born in NZ
    - permanent residents
    - those with visas that allow them to reside, work or study in NZ (including international students and temporary workers)
    - those living and working here without a requirement for a formal visa (e.g. Australians living in NZ)
- It excludes short term visitors (such as tourists)

Data in the spine

- Births (Department of Internal Affairs)
- Visa data (Ministry of Business, Innovation and Employment) – excluding visitor and transit visas
- Tax data (Inland Revenue)

IDI spine




DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
CHARGES
SENTENCING
VICTIMS
POLICE
JUSTICE
CORRECTIONS
Privacy by design
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

How is the data kept safe?

We operate a 'five safes' framework to ensure that access to the IDI and LBD is only provided if all of the following conditions can be met:

Tikanga-based framework in alignment with IDI 5 Safes

Safe people	Pūkenga (Skill, Expertise)	Whakapapa (Relationships)
Researchers can be trusted to use data appropriately	Researchers have experience researching Te Ao Māori topics	Researchers have existing relationships with the communities the data originates from
Safe Projects	**Pono (Truth, Validity)**	**Tika (Correct, Accuracy, Fairness)**
Project has a statistical purpose and is in the public interest	Level of accountability to community of research is explained	Use of data will add value and improve outcomes for Māori and NZ
Safe Settings	**Kaitiaki (Guardians)**	**Wānanga (Repositories of knowledge)**
Preventing unauthorised access to data and ensuring data is secure	Decision-makers of the project are identified and Māori are involved in decision-making	Institutions have established systems, policies and procedures to ensure data is used in culturally appropriate and ethical ways
Safe Data	**Wairua (Spirit and essence of people)**	**Mauri (Life force principle)**
Personal information is not identified	Māori community objectives align with project research objectives	Level of transformation of the data from its original collection purpose is explained
Safe Output	**Noa (Ordinary, Unrestricted)**	**Tapu (Prohibited, Restricted, High sensitivity)**
Outputs are confidentialised and individuals are not identified	Accessibility of data and awareness of the impact on Māori community	Sensitivities in the use of data are identified including privacy issues for individuals and communities

Access to the IDI for research

Go

You need five greens before microdata access can be considered

- 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

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

- Research is for commercial gain
- Unrecognised research institution
- Data integration is required
- Researchers are based overseas
- Quality of data is not adequate for research purpose
- Lead researcher is a student or junior staff member
- Research is for service provider evaluation
- Non-government researchers are requesting access to business tax data

Stop

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

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

Before applying

- Website: stats.govt.nz/idi
- Talk to us
- Researchers must be: Capable, Trustworthy and in NZ
- Research must have: strong public good element

Your application

- Talk to us
- Health projects need to check if ethics approval needed
- Some data sets accesses need agency review

In the data lab

- You'll sign a researcher undertaking, declaration of secrecy and contract
- New researchers start at Statistics NZ Data Lab (confidentiality training)
- Once approved your IDI user profile is built and accesses allocated
- Output is checked to ensure correct confidentiality rules have been applied

Support

- The wiki within the secure Data Lab environment
- Online researcher and agency forum 'MeetaData
- Ongoing forums. You are encouraged to be part of the community

DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
CHARGES
SENTENCING
VICTIMS
POLICE
JUSTICE
CORRECTIONS
Where to next?
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

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

DEPARTURES
ARRIVALS
CUSTOMS
BIRTHS
DEATHS
MARRIAGES
DIA
POLICE
CHARGES
SENTENCING
VICTIMS
JUSTICE
CORRECTIONS
Questions?
access2microdata@stats.govt.nz
MIN HEALTH
ACC
MBIE
STATISTICS
IR TAX DATA
INDIVIDUAL TAX

https://www.stats.govt.nz/about-us/data-leadership