THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Data Governance

Andrew Stott
formerly UK CIO & Director, *data.gov.uk*
Senior Consultant, World Bank

Azerbaijan, 7 April 2022
v0.3

@dirdigeng
andrew.stott@dirdigeng.com

What is Data Governance about?

Data governance is the specification of decision rights and an accountability framework to ensure the appropriate behavior in the valuation, creation, consumption and control of data and analytics. – *Gartner Group*

- Data is a key asset. It needs to be:
  - Safeguarded
  - Used to maximise its value and not misused
- Data Governance is the policies & controls to do this
- Data Governance encompasses Data Management, Open Data and Data Sharing
- Data Governance is about all data, not just statistics

Objectives of Data Governance

- Maximum value from data
- More beneficial use and re-use of data
- Lower costs of data collection
- Lower frictional costs of using data
- Protection of stakeholder interests
- Transparency and accountability for collection, holding and use of data

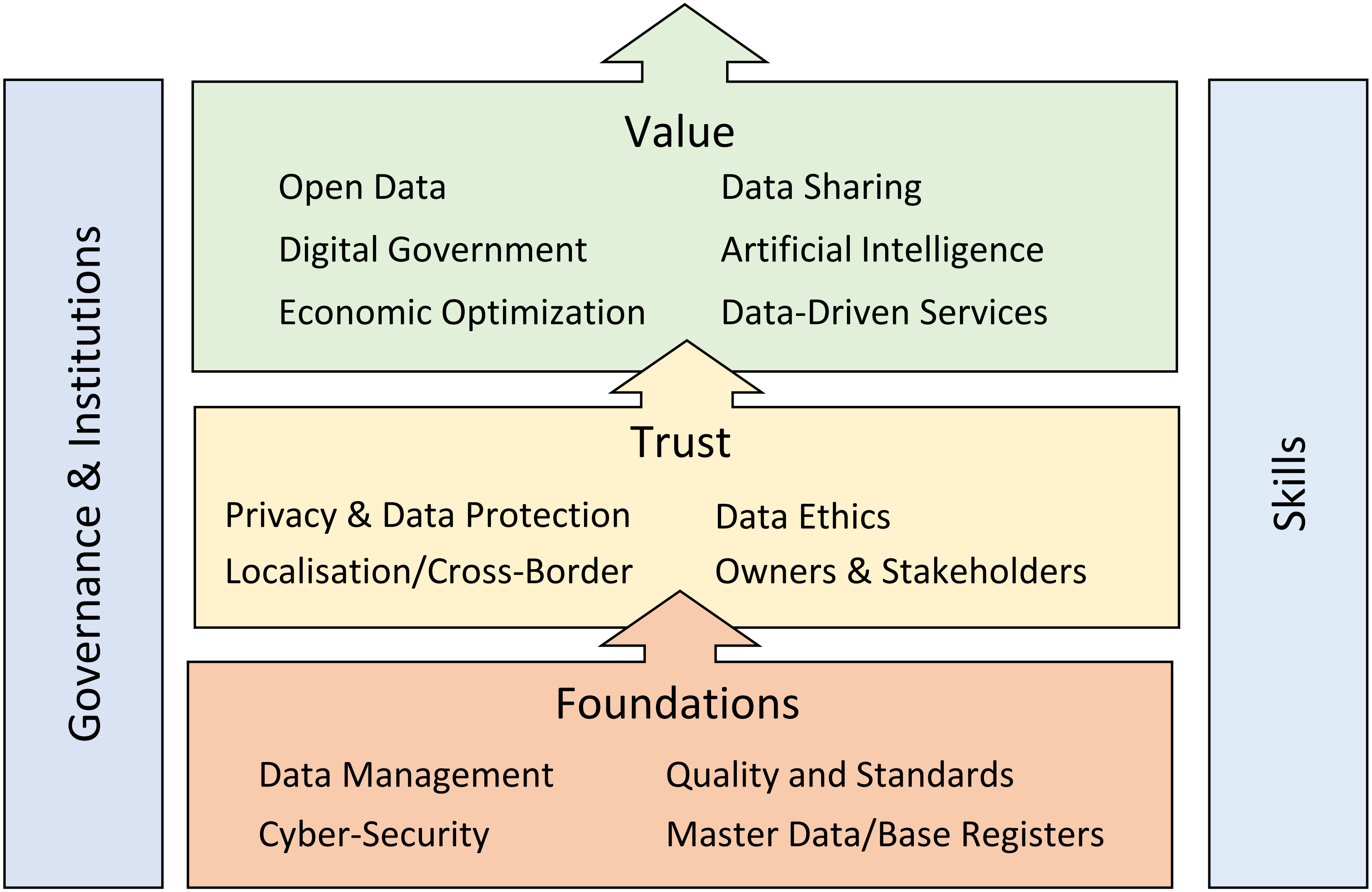
How value arises from data

Public Services	• Improved service delivery • Prioritisation of resources • Accountability
Production of Goods & Services	• Quality improvement • Cost reduction • Innovation of new lines
Optimisation of Economies	• Matching supply & demand • Lower barriers to entry • Lower transaction costs
Data-driven businesses	• Platforms • Data intermediaries • Skills-based services

Data is central to Digital Government

- Digital Government is about more than channel shift for customer contacts
- Data, accessed through secure channels from authoritative sources, is better than documents
- Use public sector and private sector data for risk-based decision making and service optimisation
- Redesigned data-driven services maximise benefits to citizens, business and government

Data Governance Framework



Foundations

Data Management

- Know what data an agency has, and make sure that it can be found and used
  - "external" metadata – what a dataset is (eg DC)
  - "internal" metadata – how to use a dataset (eg DDI)
  - Manage the life-cycle of each dataset – creation/acquisition, storage, use, disposal
- Manage the associated rights and interests in the data
- Ensure clear accountability for the data throughout its lifecycle

Cybersecurity

- Protect all three dimensions
  - Confidentiality
  - Integrity
  - Availability
- Consider cybersecurity broadly
  - Protect all interests in the data, not just owner interests
  - Protect data when others are using it
  - Counter supply chain vulnerabilities

Data Quality and Standards

- “Quality” means fits for purpose, not necessarily perfect
- A quality strategy should include:
  - Know the quality of the data (timeliness, accuracy, provenance, verification)
  - Ensure data users understand quality risks and avoid/mitigate them
  - Robust processes for continuous quality assessment and correction
  - Continuous improvement

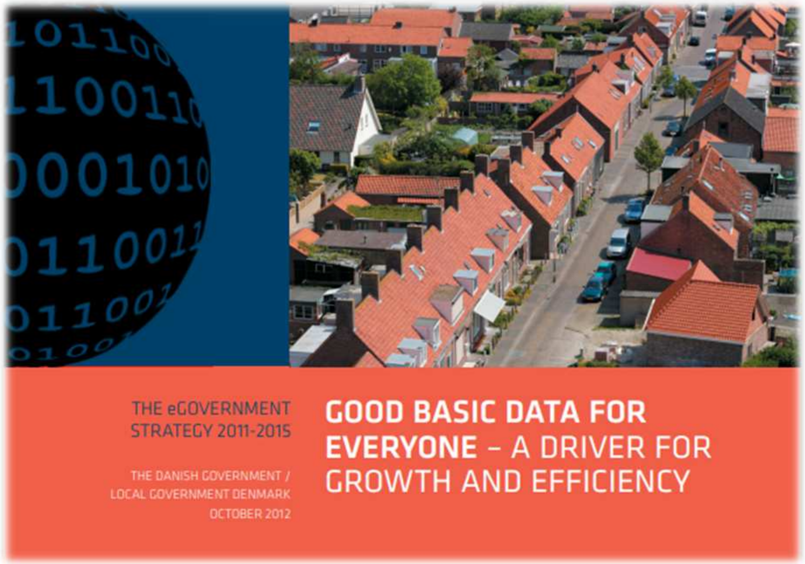
Data Standards

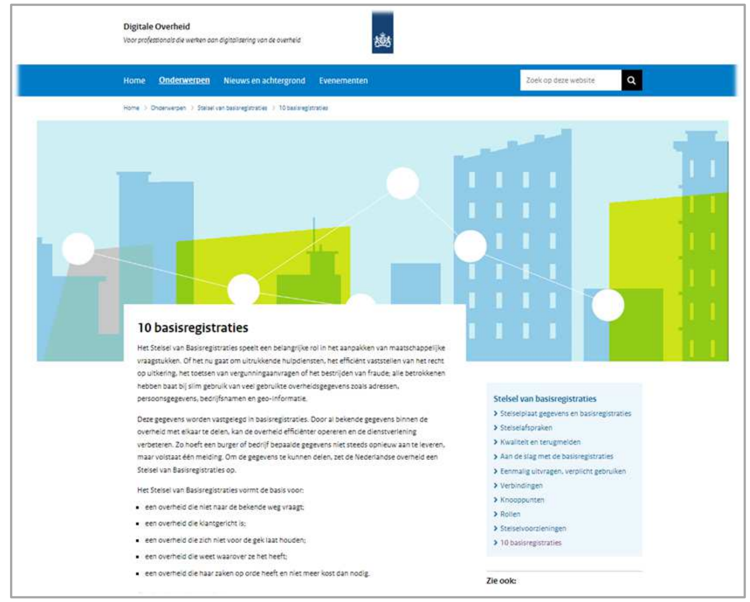
- Standards should include
  - Technical Standards
  - Interchange Standards
  - Business Semantics
  - Associated data management practices
- Standards necessary for inter-operability – but not sufficient
- An e-Government Interchange Framework (e-GIF) needs to be more than a list of ISO standards!

Master Data, Base Registers and De-duplication

- General principle of create once, use multiple times
- Use APIs, message switches (eg X-Road) to obtain information secure when needed from authoritative source – do not take copies
- “Owning” agency must take cross-Government accountability
- Digital government leaders increasingly prioritising “Base Registers”

Base Registers





- Typically 10-15 databases with wide utility
- Clear accountability for maintenance of each database for all users
- All Agencies must use Base Register rather than collect information again

Questions?

Trust

Privacy and Data Protection

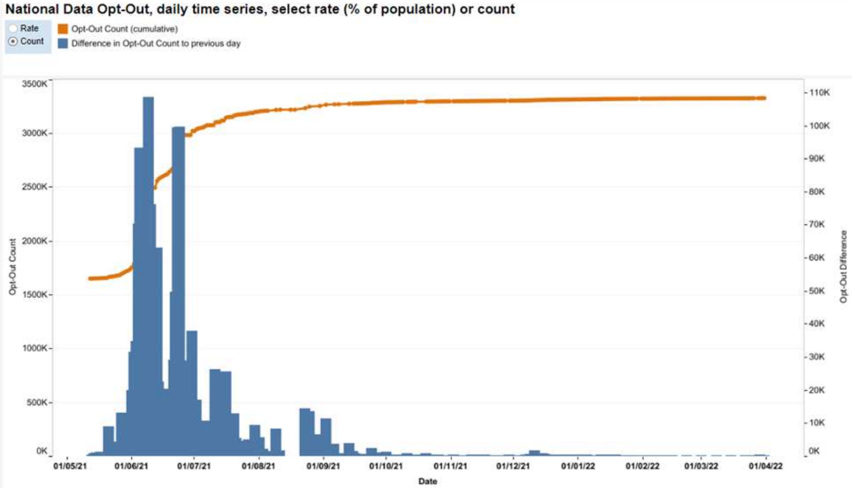
- WDR2021 recommends rights-based approach
  - Rights of data subjects as well as obligations on data users
  - Rebalance asymmetries between data holders and data subjects
  - Effective rights first, to give framework for permissible data sharing and re-use
- Use, whether by consent or otherwise, should still be necessary, proportionate and transparent
- Finessing concerns can be counter-productive

UK National Health Service data opt-outs

NHS National Data Opt-Out

Latest Opt-Out Count on 31/03/2022	Latest Opt-Out Rate on 31/03/2022
3,328,303	**5.41%**

National Data Opt-Out, daily time series, select rate (% of population) or count



The Guardian
UK edition
For 200 years
News website of the year

NHS data grab on hold as millions opt out

A plan to share GP data was set to launch in September, but an online summer campaign has prompted widespread dissent

More than a million people opted out of NHS data-sharing in one month in a huge backlash against government plans to make patient data available to private companies, the *Observer* can reveal.

But after the proposals were quietly announced in May, doctors' leaders objected to the short six-week deadline for the public to opt out of the scheme, while privacy campaigners warned the process to remove identities could be reversed.

The deadline was initially delayed to September, but an online campaign encouraging people to opt out grew over the summer. Government figures show that in May 107,429 people opted out. In June, a further 1,275,153 followed.

Data Ethics

- "Can I? Should I?" – *Winston Churchill*
- A Data Ethics Framework should include
  - Transparency about data usage, benefits, risks and processes
  - Accountability through effective governance and oversight, and involving diverse expertise
  - Fairness through ensuring just and non-discriminatory outcomes
  - Restraint through recognising and respecting the quality and limitations of the data
- Wide consultation about novel data usage

Algorithmic decision-making

- A lot of active policy development currently
- Emerging expectations
  - Evaluation for hidden biases in algorithm and training data
  - Transparency that an algorithm is being used
  - Algorithm’s result should be explained
  - Human takes final decision for most important decisions (eg medical)
  - Right to have decision reviewed
- Important to track developments in EU and elsewhere

Localization & Cross Border Data Flows

- Ensuring a country can benefit from its own data
- Ensuring protections are preserved as data is moved across borders
- Ensuring that essential data-driven processes are resilient

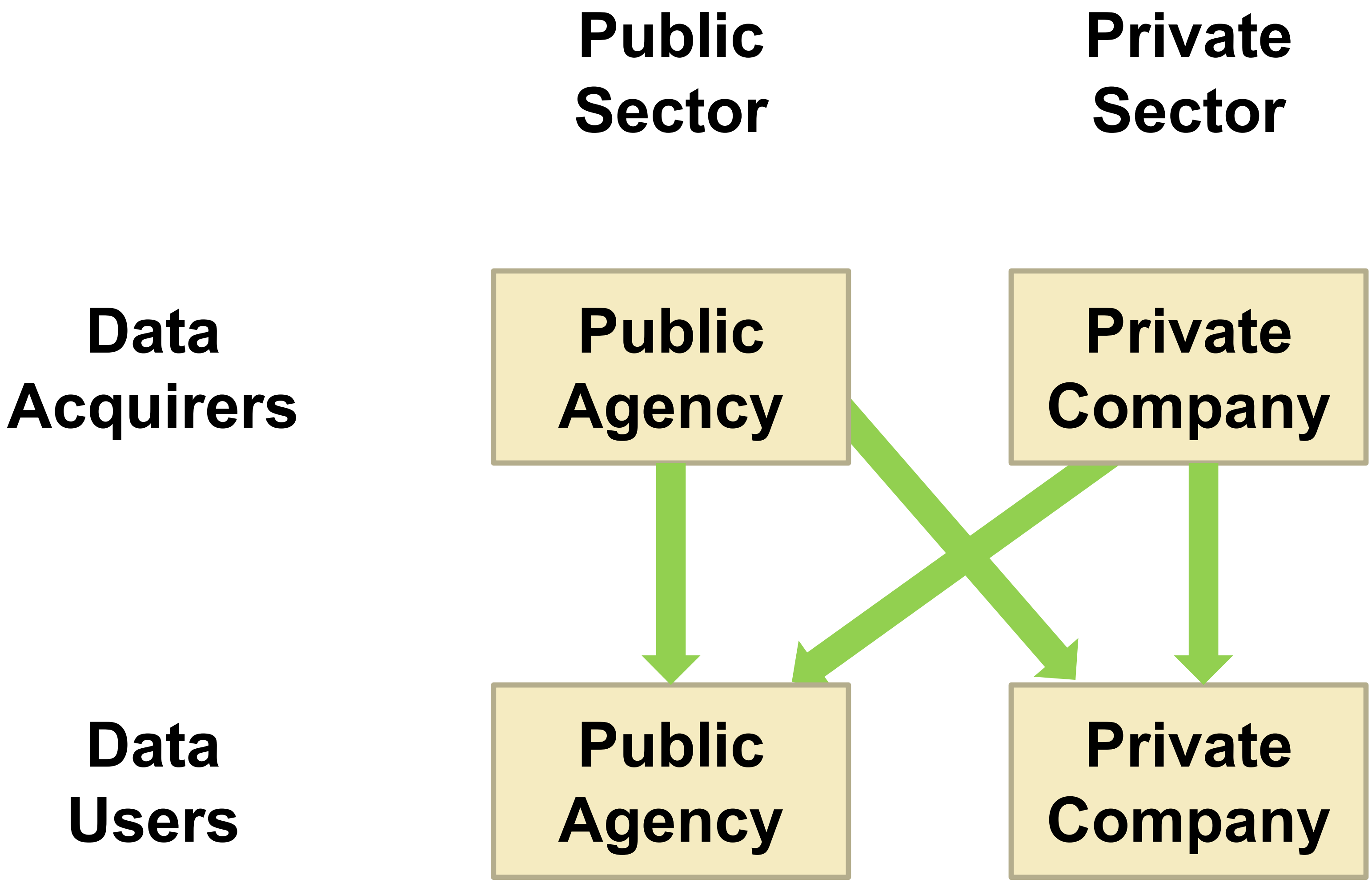
Ownership and Stakeholders

- “Ownership” not as simple as for physical assets
- Stakeholders may have ownership-like rights too
- Issues need to be resolved to allow data to be used for maximum value

Questions?

Value

More value if users can access more data



(Not shown: academic, citizens, …)

Mechanisms for accessing more data

Holder	User	Key Policy Mechanisms
Public agency	Public agency	• Data sharing policy • Catalogues and metadata • Open Data policy • Key registers and master data
Private company	Public agency	• "Public service" terms • Fair Regulatory requirements
Public agency	Private company	• Open Data policy • Open Metadata policy • "Walled Gardens" • Industry Pools
Private company	Private company	• Data Pools • Data Exchanges • Metadata transparency • Patent-style terms

Governance and Institutions

Center	• "Chief Data Officer"? • Central Data Governance & Policy Team • Standards-setter for data and processes
Whole of Government	• "Data Council"? • Data Governance and accountable post in each agency's accountability mechanism • Cross-agency governance of each Key Register
Independent/ External	• Privacy/Data Protection Commissioner • Data Ethics Panel • Sector Panels • Business Data Promotion Body

Getting Started

A organisation that thinks it does not have a data governance problem ….

… has a problem!

Getting Started

Strategic Issues

- Multi-year program
- Many actors
- Regulation alone not enough
- Cultural issues
- Network benefits

Key Change Management Approaches

- Clear Vision
- Prioritise
- Small steps
- Show benefits
- Form teams
- Build skills
- Celebrate Success
- Communicate

Getting Started

- Initial stocktake/benchmark with key stakeholders
- Raise awareness and share national and international good practice with stakeholders
- Set strategic vision and potential benefits
- Establish leadership
- Form central team and “Data Council”
- Quick wins by ad hoc initiatives and chosen sectors to build skills and credibility
- Prioritise by benefits, needs of digital government and sectors

How the World Bank can help

- Experience-sharing from experts and leading countries in the field of Data Governance
- Analytical reports and guidance
- “Data Ecosystem Maturity Assessment” *(under development)*
- Expert assistance with initial benchmarking, strategy development and program design
- “Convening power” with external stakeholders as part of awareness-raising

Questions and Discussion

Case Study

Data for Digital Government

(if time allows)

UK annual car tax – old process

- Assemble Documentation
  - Ownership Certificate
  - Insurance Certificate
  - Safety Certificate
- Fill in Form
- Take all four documents to a Post Office
- Queue
- Pay money
- Be issued with Tax Disc to stick on windscreen

UK annual car tax – old process

Disadvantages

- Costs for government, citizens and businesses
- Errors in recording
- Genuine documents (certificates, tax discs) were attractive to criminals to sell stolen cars (2 inspectors were murdered)
- Documents were forgeable
- Police database ~6 weeks out of date
- Enforcement of requirements only once a year

UK annual car tax – new process

Data-driven approach

- Industry-run up-to-date insurance database
- Government-run database of ownership
- Private repair shops doing safety checks update Government-run database
- Online transaction checks all three databases in real time
- Police database updated within 24 hours
- Tax status can be checked online – no disc

UK annual car tax – new process

Benefits to all stakeholders

- Citizen: takes 3 minutes a year, available 24/7, and no certificates to keep/lose
- Taxpayer: substantial cost savings in transaction costs and records, and enforcement
- Crime prevention: no valuable documents, up to date police records
- Policy effectiveness: continuous enforcement, including enforcement of insurance

UK annual car tax – new enforcement