ODI Leaning in Tanzania

Dr David Tarrant | @davetaz | The Open Data Institute

http://tanzania.learndata.info

Content created by
The Open Data Institute

The characteristics of open data

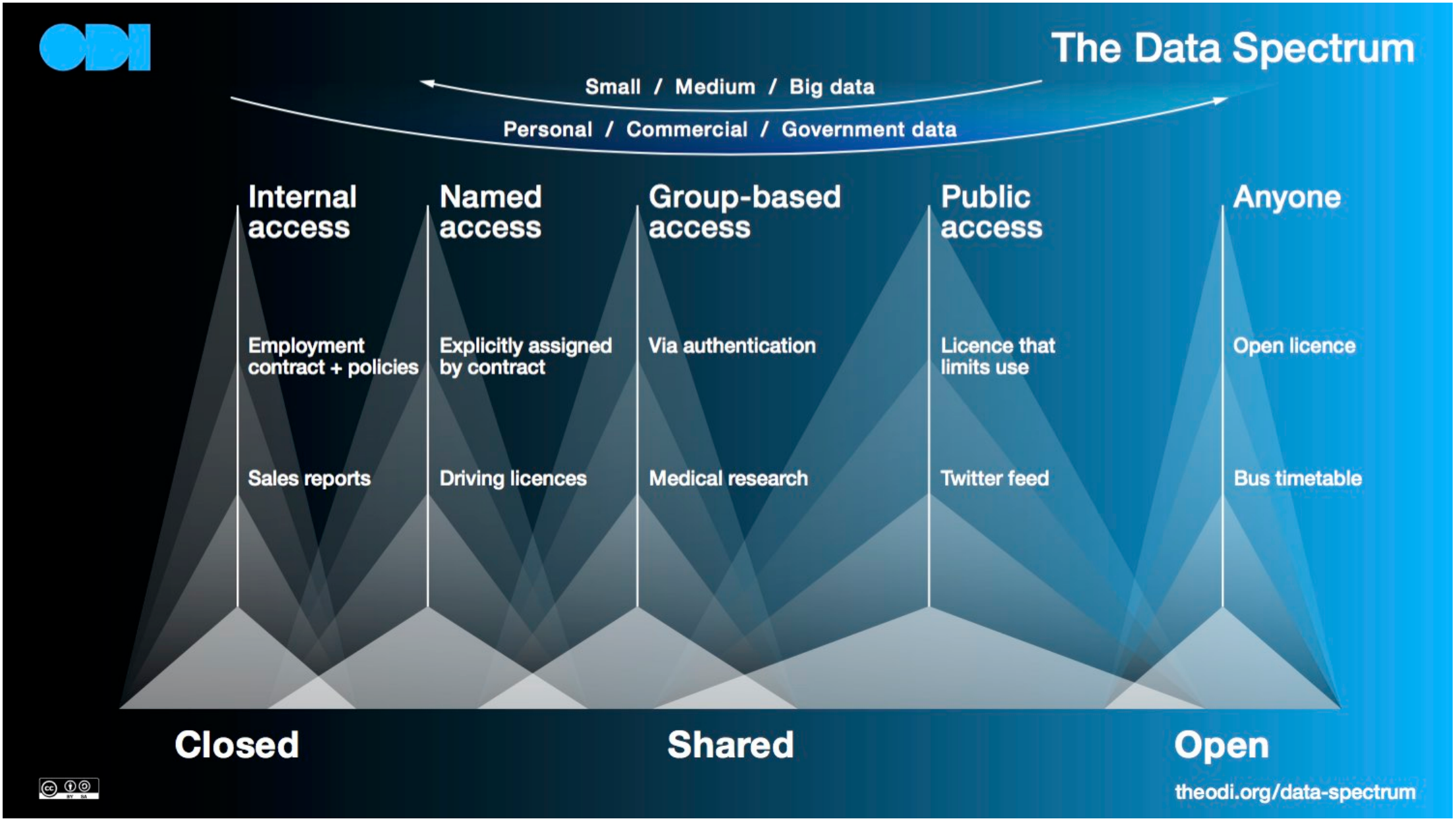
Content created by
The Open Data Institute

Outcomes

Identify a number of different characteristics of data

Explain the justifications for publishing different types of data

Evaluate the current open data ecosystem and future opportunities


The Data Spectrum
Small / Medium / Big data
Personal / Commercial / Government data
Internal access
Named access
Group-based access
Public access
Anyone
Employment contract + policies
Explicitly assigned by contract
Via authentication
Licence that limits use
Open licence
Sales reports
Driving licences
Medical research
Twitter feed
Bus timetable
Closed
Shared
Open
theodi.org/data-spectrum

Exercise (part 1)

Pick three datasets, open, shared and closed from the spectrum.

Write the dataset title on a post-it

Write open, shared or closed on the post-it

Types of Data

Reference data

“things”

People Facilities Places
Books Buildings

Transaction data

“stats involving things”

Expenditure Weather
Consumption

Observation

Exercise

Categorize your data into reference and transactional data.

If they are all in one category you have 2 minutes to add some new datasets to the empty category.

When done, put a “T” or and “R” on each dataset post-it.

Update frequency

Static | In frequent updates | Frequent updates | Live

Exercise

Categorize your data into **frequency of updates**

If they are all in one category you have 2 minutes to add some new datasets to the empty category/ies

Put a number on your post-its representing the frequency of updates.

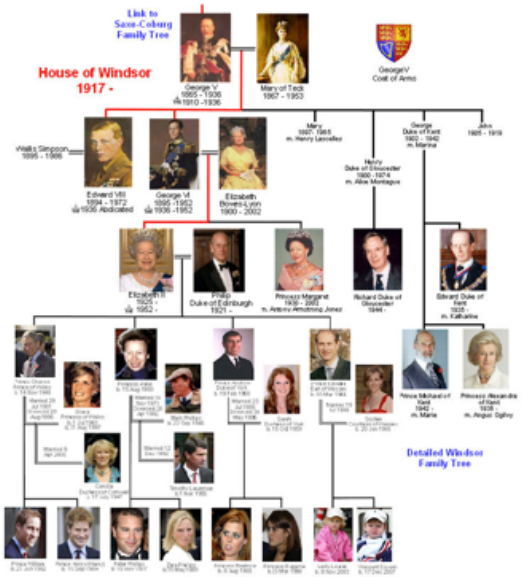
0 = static, 1= In frequent, 2 = Frequent, 3 = Live

Data Representations

Tabular

Region	Production		
Country Level 2	Production (thousand MT)	Change from last year	Change from 5 year average
Brazil	57289	-4.05%	+2.68%
Mato Grosso	18,008	0.90%	6.17%
Parana	9,571	-19.55%	-9.08%
Rio Grande do Sul	7,844	0.88%	9.35%
Goias	6,820	4.23%	5.27%
Mato Grosso do Sul	4,218	-7.69%	-1.97%
Minas Gerais	2,667	5.12%	2.41%
Bahia	2,512	-8.58%	4.84%
Sao Paulo	1,392	-3.77%	-6.81%
Maranhao	1,087	-13.93%	0.58%
Santa Catarina	1,039	9.81%	13.35%
Tocantins	902	-0.90%	7.05%
Piaui	856	4.49%	23.75%
Para	194	-3.13%	1.02%
Distrito Federal	155	1.37%	-1.11%
Roraima	22	-54.10%	-41.70%

Hierarchical



Network/Graph

Exercise

Categorize your data into **tabular, hierarchical (tree) and graph (network)**

If they are all in one category you have 2 minutes to add some new datasets to the empty category.

Add the word “**tab**”, “**tree**” or “**net**” to your post-its to represent the different structures.

Justifications

Trust and Transparency

Enabling the economy

One more

Categorize your data into **transparent** and **enabling**.

Summing up

Do you have any obvious grouping of your datasets?

Is this reflective of the whole open data ecosystem?