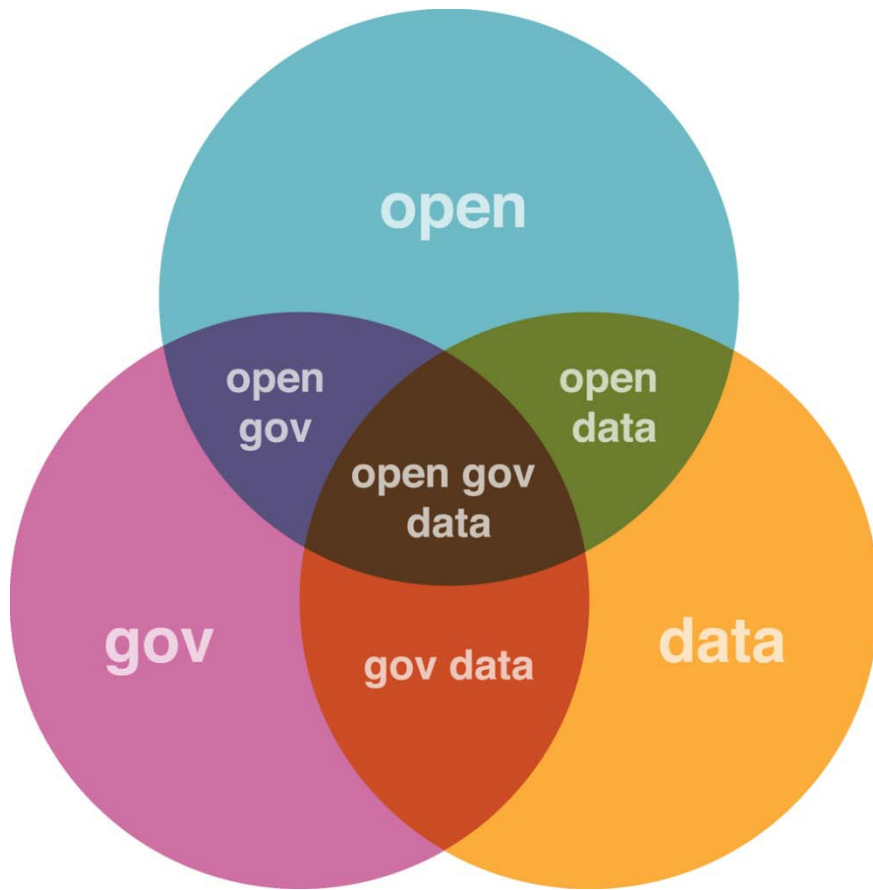


# Creating the right environment to support safe access to microdata

Tenth Management Seminar for the  
Heads of National Statistical Offices in Asia and the Pacific  
7 – 8 December 2011, Chiba, Japan

# The era of Open Data

- ▶ “Open” is now a key word for statisticians
- ▶ “Open data” for:
  - **Open government**: give access to data to your citizens for participation and transparency
  - **Open knowledge**: give access to your data to researchers
    - Create incentives for intermediaries to use and re-purpose your data
    - Because others may do it better, cheaper, differently ...



**Open Government** is about making core government information and processes more accessible to citizens for transparency, participation and innovation.

**Open Data** is a term applied to any data that is *legally and technically open*: free to use with little restriction and available online in standard, re-usable file formats.

**Open Government Data** sees open data principles applied to any government produced data such as budgets and finances, statistical data, administrative data, public service and infrastructure data.

# Some examples ... World Bank

**THE WORLD BANK**  
Working for a World Free of Poverty

ABOUT DATA RESEARCH LEARNING NEWS PROJECTS & OPERATIONS PUBLICATIONS COUNTRIES TOPICS

## Data

By Country By Topic Indicators Data Catalog Microdata News About For Developers Products

This page in [English](#) | [Español](#) | [Français](#) | [العربية](#) | [中文](#)

### New data and tools on climate change

A new Climate Change Knowledge Portal, launched today, includes visualization tools depicting temperature and rainfall scenarios to the year 2100 and the Apps for Climate competition will kick off in December at the United Nations Climate Change Conference in Durban, South Africa.

**THE WORLD BANK**  
**Open Data**

*The Data Catalog provides download access to over 7,000 indicators from World Bank data sets.*

**Find an indicator**  
GNI per capita, Atlas method (current US\$)    
[Learn more about data sets](#)

**BROWSE DATA**

- By Country**  
View profiles for over 200 countries and economies
- By Topic**  
See key indicators for a variety of topics
- Indicators**  
View data for over 1,200 indicators
- Use our Data**  
Learn how to use World bank's data

<http://data.worldbank.org/>

# Some examples ... UNDP

The screenshot shows the UNDP website's 'Open Data' section. The header includes the UNDP logo and navigation links: 'UNDP around the world', 'Operations', 'Research & Publications', and 'News Center'. Below the header, there are three main navigation buttons: 'Our Work', 'Millenium Development Goals', and 'Our Perspective'. The breadcrumb trail reads 'Home > Operations > Transparency > Open Data'. The main content area is titled 'Open Data' (circled in red) and 'UNDP Supports Projects'. A pie chart displays the distribution of projects by country/region. The largest slice is 'Other', followed by 'Afghanistan'. Other countries include Sudan, Zimbabwe, Democratic Republic of Congo, Bureau for Dev. Pol., Brazil, Bangladesh, Pakistan, Prog for Palestinia..., Egypt, Honduras, Guatemala, Indonesia, United Republic of Tanzania, Tajikistan, and Chad. A pagination indicator shows '1/3' with navigation arrows.

Home > Operations > Transparency > Open Data

Overview

UNDP and Transparency

Exploring the Data

- By Country/Unit
- By Project

**Open Data**

**UNDP Supports Projects**

1/3 | ◀ | ▶

Afghanistan

Sudan, Republic of ...

Zimbabwe

Democratic Republic...

Bureau for Dev. Pol...

Brazil

Bangladesh

Pakistan

Prog for Palestinia...

Egypt

Honduras

Guatemala

Indonesia

United Republic of ...

Tajikistan

Chad

Other



# Some examples ... US Govt

An Official Web Site of the United States Government Wednesday, November 23, 2011 Text: A+ A- A Share



HOME DATA ▾ APPS COMMUNITY METRICS OPEN DATA SITES GALLERY WHAT'S NEW



## ENVIRONMENTAL COMPLIANCE AND ENFORCEMENT DATA

Enforcement & Compliance History Online (ECHO) provides information about compliance with environmental laws that ensure clean air, water, and land. Use interactive menus and maps to search for facilities that were inspected, who had violations, whether EPA or your state took enforcement actions and more.



**ECHO**  
Enforcement and Compliance History Online

### Latest Datasets

- Metropolitan Area Look-Up
- FY11\_EOM\_Sep\_VGLI Coverage Amount by State...
- Number of Veterans Insured by VGLI by State...
- FY10\_Life Insurance Payments, Face Value of...
- Total Number of Railroad Employees by State...
- Distribution and Production of Oil and Gas...
- Monthly Utility and Nonutility Fuel Receipts...
- Monthly Energy Consumption by Sector
- Monthly Natural Gas Consumption by End Use
- Annual U.S. Natural Gas Summary

### DATA AND APPS



### COMMUNITIES



### OPEN GOVERNMENT



# Some examples ... UK Govt

The screenshot shows the data.gov.uk website. At the top left is the HM Government logo. The main header features the text "data.gov.uk BETA" and "Opening up government", which is circled in red. To the right is a search bar with the placeholder text "What are you looking for?" and a green "Search" button. Below the header is a navigation menu with links for Home, Data, Apps, Event, Forum, Blogs, Consultation, Linked Data, Resources, and About. In the center, there is a search box with a "Go" button and the text "Enter your postcode to find tools and information about your local area using government data". Below this are six featured content cards, each with a title and a representative image: "Who does what in Whitehall?" (meeting room), "How your money is spent" (stacks of coins), "Government contracts in full" (hand writing on a document), "Who ministers are meeting" (group of people), "Business plans" (business meeting), and "Applications" (a mobile app interface for "Crime Spy UK" showing location "Cambridge City West (cambridgeshire)").

<http://data.gov.uk/>

# Some examples ... Kenya

The screenshot shows the Kenya Open Data website. At the top, there is a navigation bar with links for Home, Data Catalog, Suggest a Dataset, Community Applications, and Counties. On the right side of the navigation bar, there are links for Sign Up and Sign In. The main header features the Kenya Open Data logo, which consists of a red square with a white 'K' and the text 'KENYA openData' next to it. Below the logo, there is a section titled 'Open Data: The bigger picture' with a paragraph of text and a link to 'Read more about this site here.' To the right of this section is a green button that says '+ Developers Click Here'. Below the 'Open Data' section is a section titled 'Counties' with a sub-heading 'Select a county to learn more about that county:' and a map of Kenya. The map shows three counties highlighted: Turkana (green), Mandera (yellow), and another county (pink). To the right of the map is a sidebar with three sections: 'PROJECT TOPICS' with a list of topics (Education, Energy, Health, Population, Poverty, Water & Sanitation), 'PUBLIC EXPENDITURES' with two links ('View Public Expenditures »' and 'View CDF Spending »'), and 'DEVELOPMENT INDICATORS' with one link ('View World Development Indicators »').

Home Data Catalog Suggest a Dataset Community Applications Counties Sign Up Sign In

**KENYA**  
openData

### Open Data: The bigger picture

This site makes public government data accessible to the people of Kenya. High quality national census data, government expenditure, parliamentary proceedings and public service locations are just a taste of what's to come. There's something for everyone: maps to start exploring, interactive charts and tables for a deeper understanding, and raw data for technical users to build their own apps and analyses. Our information is a national asset, and it's time it was shared: this data is key to improving transparency; unlocking social and economic value; and building Government 2.0 in Kenya.

[Read more about this site here.](#)

**+ Developers Click Here**

### COUNTIES

Select a county to learn more about that county:

Turkana Mandera

### PROJECT TOPICS

- Education
- Energy
- Health
- Population
- Poverty
- Water & Sanitation

### PUBLIC EXPENDITURES

- [View Public Expenditures »](#)
- [View CDF Spending »](#)

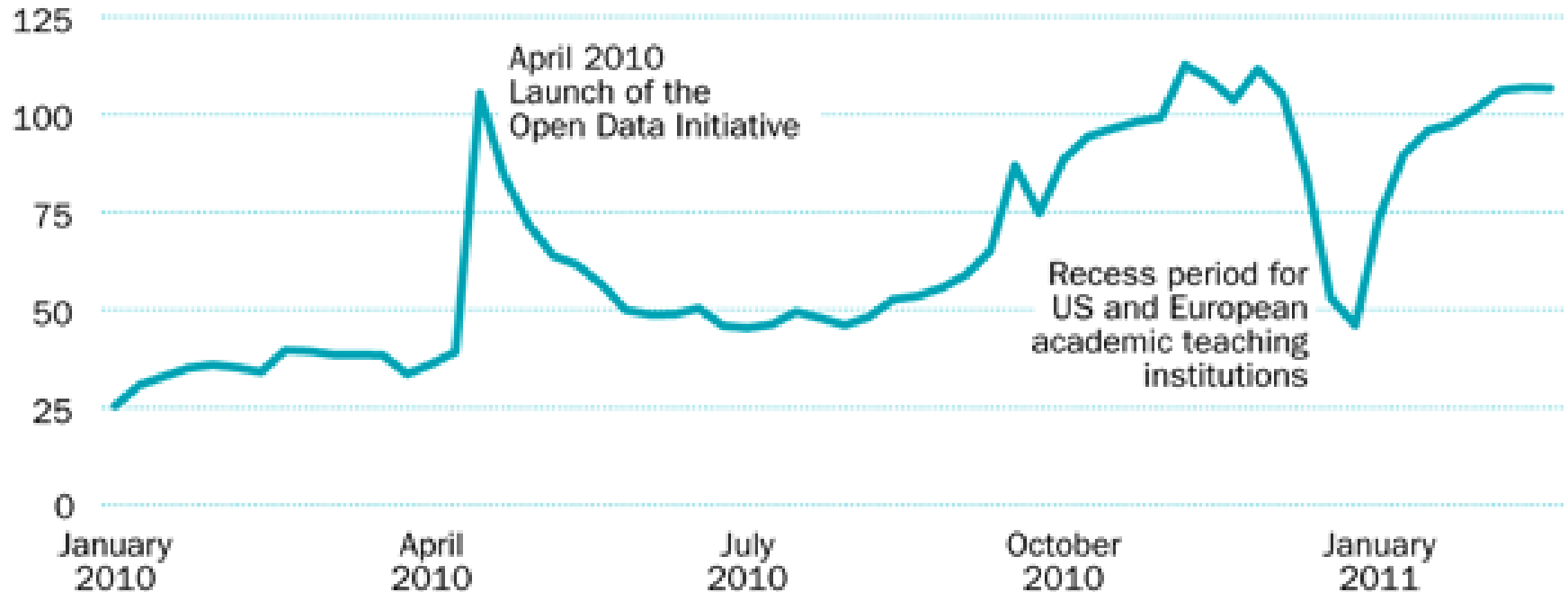
### DEVELOPMENT INDICATORS

- [View World Development Indicators »](#)



# Openness increases demand

Weekly unique visitors to <http://data.worldbank.org> (thousands)



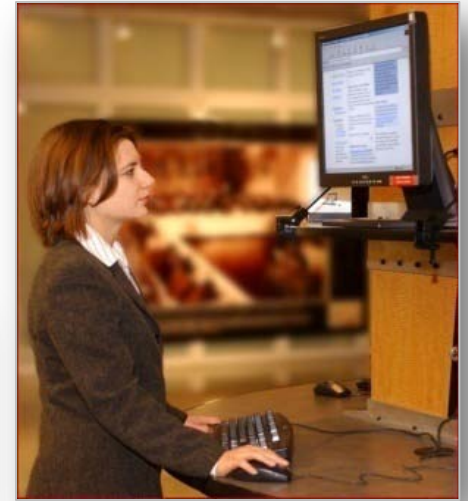
Source: World Bank staff calculations from Omniture data.

World Bank Open Data Initiative  
<http://data.worldbank.org>  
→ number of visitors multiplied by 4

# High expectations

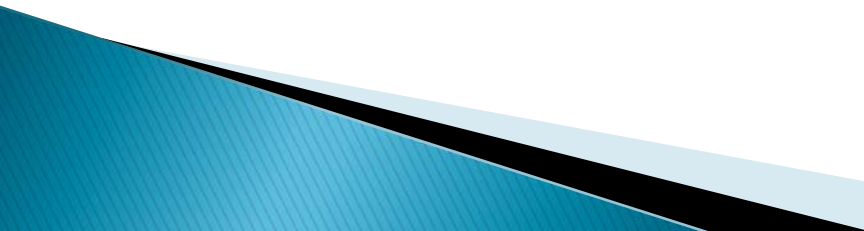
- ▶ High expectations on statistical agencies
  - From sponsoring agencies who want data to be used
  - From users who need more diverse and disaggregated data
  - From the citizen who expects transparency
- ▶ Need to “liberate” microdata
  - Not a new movement, but not (yet) generalized
  - E.g.: Statistics Canada launched its Data Liberation Initiative in 1996 ; India has a long tradition of sharing microdata

# Need to modernize infrastructure



# Infrastructure for microdata

Statistical agencies need to build a proper infrastructure for managing and disseminating microdata. Infrastructure involves:

- ▶ IT infrastructure
  - ▶ Standards
  - ▶ Corporate policy and procedures
  - ▶ Legislation
  - ▶ Trained staff
  - ▶ Sustainable funding
- 

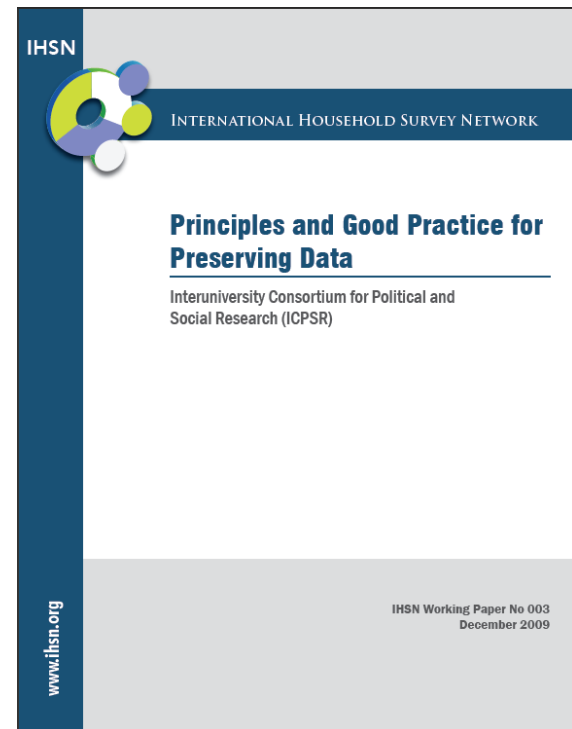
# IT infrastructure

- ▶ File servers and web servers
  - Microdata are still often stored on individual hard drives → loss of data, multiple versions, etc.
- ▶ Data security: need a strong enforcement of user authentication system, and antivirus
- ▶ Backup system (local and in remote location)
- ▶ Regular migration of formats and media
- ▶ Software
  - Data entry, editing, analysis
  - Data documentation and dissemination



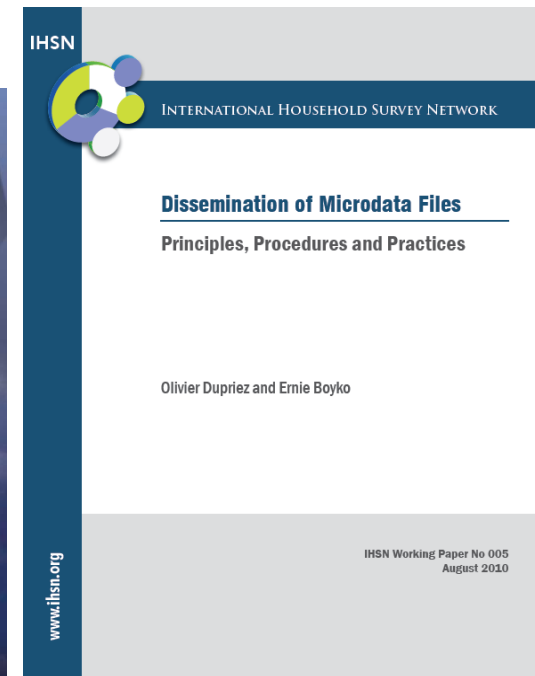
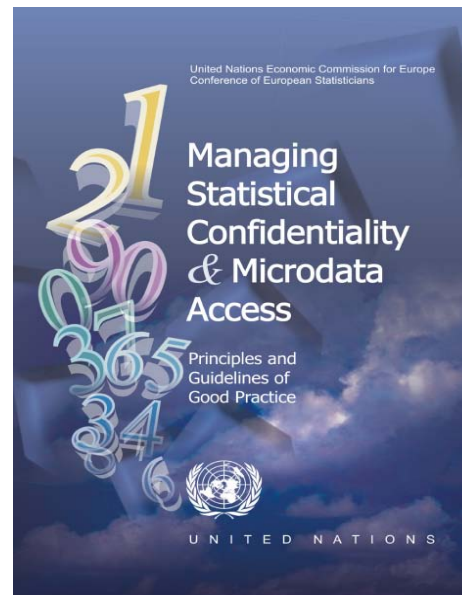
# Standards

- ▶ Compliance with international standards fosters quality, efficiency and effectiveness and collaboration with other agencies
- ▶ Metadata standards:
  - **Data Documentation Initiative (DDI) !**
- ▶ Preservation standards
  - Open Archival Information System Reference Model (OAIS)



# Corporate policy & procedures

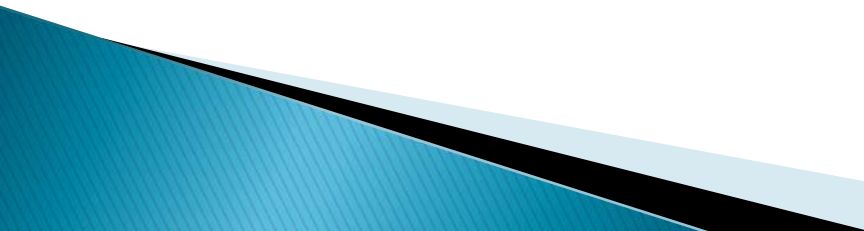
- ▶ Publish a formal dissemination policy and protocols, with:
  - Statement on privacy and confidentiality protection
  - Data access procedure
  - Clear terms of use
  - Pricing policy



<http://www.unece.org/fileadmin/DAM/stats/publications/Managing.statistical.confidentiality.and.microdata.access.pdf>

<http://www.ihsn.org/home/download.php?file=IHSN-WP005.pdf>

# Legislation

- ▶ Microdata dissemination must be in compliance with national legislation.
  - ▶ Not all microdata will be disseminated; must strike a balance between dissemination and privacy protection.
  - ▶ Increasingly, statistical legislations make (some) microdata dissemination mandatory (a good example: Mexico).
  - ▶ If legislation is outdated: change it, don't ignore it.
- 

# Trained staff

- ▶ Important to have staff trained in microdata curation, i.e. specialized in:
  - Documentation of surveys (using the DDI standard)
  - Preservation of data and metadata
  - Support to data users
  - Maintenance of survey catalogs and repository
- Important for training centers to provide such training as part of regular training curriculum and special courses for professional statisticians

# Sustainable funding

- ▶ Documenting, preserving and disseminating microdata has a cost
  - ▶ This cost is small compared with the cost of producing data, and is a good investment
  - ▶ Must be budgeted !
- 