

Open Data and Open Tools: The Global Impact Study Inventory and Web Application

Authors: Chris Rothschild, Araba Sey, Alexander Tulinsky

What is Open Access to Research Data?

Is it free and open access to raw data? Access to data collection tools and instruments? Access to data with tools to organize and analyze? These are the types of questions the Global Impact Study is considering as we develop an open approach to our research on the impacts of public access to information and communication technologies (ICTs). The approach we have adopted includes developing an open-access web application to enable meaningful sharing of the data in our inventory of public access ICT venues. Users can query, export, and map data on the location and characteristics of venues. In addition, the study is making public all source code used for the web application. This poster illustrates the types of data accessible through the application, query and visualization tools, and sample outputs. The green boxes call out some of the primary open data issues associated with the application's main features. Based on what the web application allows, to what extent do you think our data platform achieves open access goals and principles?

Global Impact Study

Goal: To investigate the impacts of public access to ICTs in venues such as libraries, telecenters, and cybercafés.

Timeline: 2007–2012

Primary Research Questions:

1. What are the social and economic impacts of public access to ICTs?
2. What is the magnitude of these impacts and how can we measure them?
3. What is the relationship between the costs and benefits of providing and using public access ICTs?

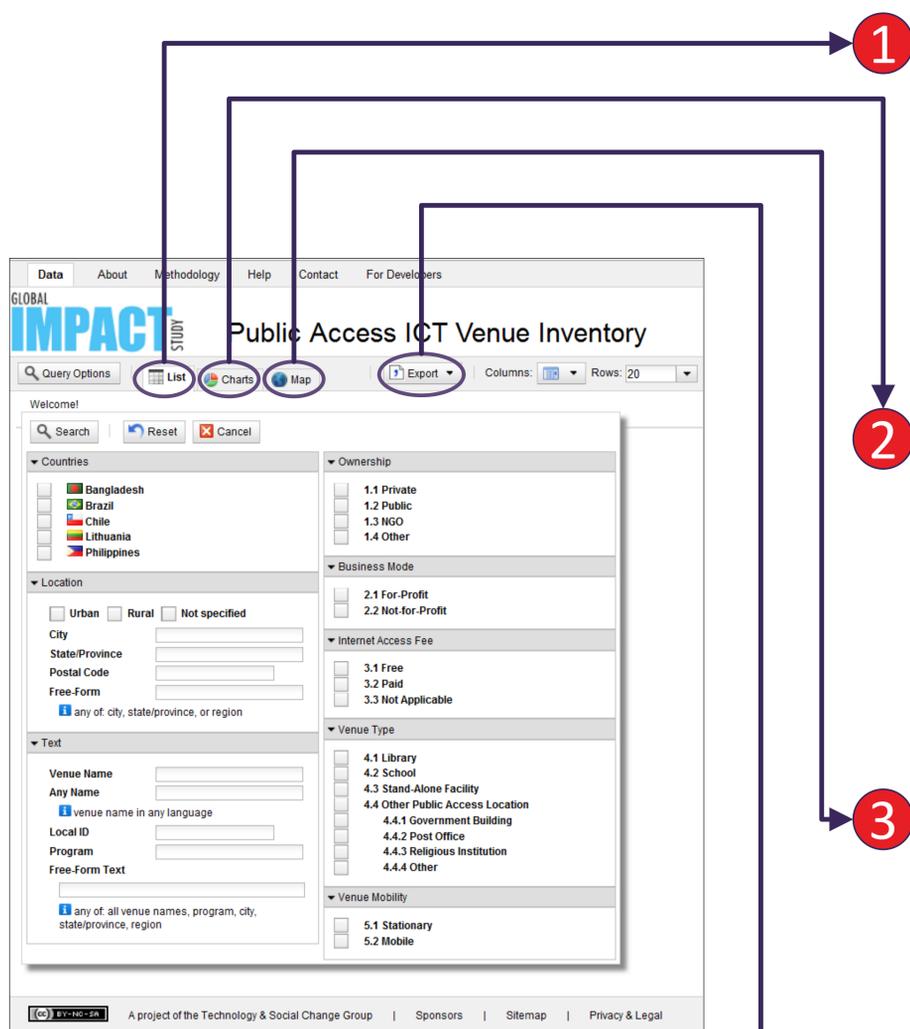
Inventory

Why: To capture the magnitude and distribution of the venues studied in the project

What: Address, venue type, and venue characteristic data

Where: Bangladesh, Brazil, Chile, Lithuania, Philippines

How: Through publically available, administrative data



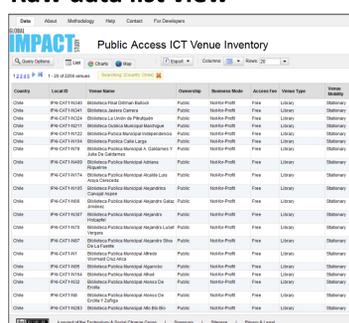
This is the home page of the application. From here users can:

- 1 Query and view data as a raw list
- 2 View the queried data as pie charts
- 3 Map queried data
- 4 Export queried data

Open data: Tapping all sources

We know that there are limitations to our data collection methods and users may find inaccuracies or gaps in some of our data. However, public users can currently only download data. How can we enable users to correct and supplement our data and at the same time maintain the integrity of the data?

Raw data list view



Selecting the "List" tab allows users to view the raw data as a list and to select query criteria.
(Left: raw data list view for "Santiago")

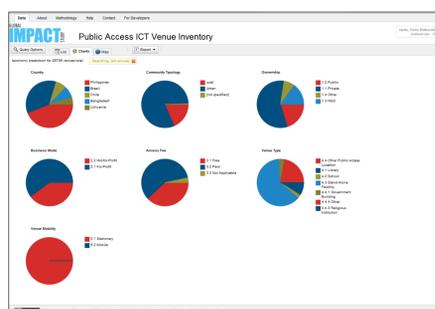
Open Access: When does public data sharing begin?

We do not release our datasets as soon as they are received.

Before we share data we want to provide time to:

1. internally verify the data (to make sure the public data are "good" and limitations have been articulated)
2. give project researchers first priority to analyze the data

Basic statistics view

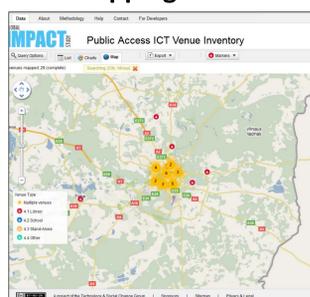


Selecting the "Charts" tab plots the queried data as a series of predefined pie charts based on venue characteristics, country, and urban/rural definition.
(Left: Pie charts compiling data for all countries)

Open Access: Adding value to open data

Providing data in a usable format is an important aspect of our stance on open data. Features like Charts and Mapping allow more meaningful sharing of data than simply posting spreadsheets.

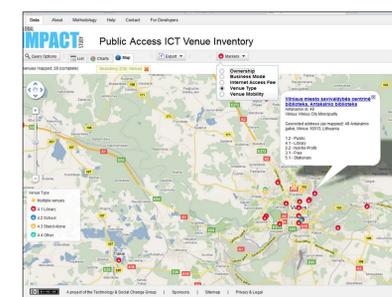
Data mapping view



Wide View

Selecting the "Map" tab plots the geocoded venues on a map. When zoomed out, venues are clustered by proximity, with the number of venues in each cluster indicated.

(Above: map displaying clustered data for Vilnius)



Zooming in

While in the map view, zooming in allows users to see markers representing the location of venues. Markers are color-coded by venue characteristic. Clicking on an individual venue shows additional venue data.

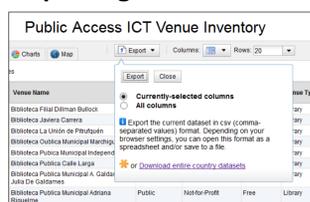
(Above: Vilnius data with a venue selected.)

Open Access: What data are "open"?

Open: Many fields, including venue characteristics; venue start, close, and future start dates; venue program affiliation; and data source.

Not open: Contact names and phone numbers. Although the data were collected from publically available sources, some contact information is kept private due to ethical and privacy concerns.

Exporting



All data-view tabs have an "Export" button that allows users to download queried data. On an additional page users can download pre-generated datasets with all the data for each country.

(Left: popup for exporting data)

Open Access: Conditions on use

Before users view and download data, they must register with personal and institutional information and agree to abide by usage conditions, including intellectual property requirements.