

citizen-generated data

Citizen-generated data is a broad category that includes any information that can be collected from people either by active involvement (experiences, ideas, upvotes) or passively (e.g. wearables or transactions data).



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: I PAID A BRIBE

I Paid A Bribe attempts to tackle corruption in India and other countries by tracking bribe payment activity and raising awareness about the nature and spread of bribe-related exchanges. The platform crowdsources anonymous reports by people who have either paid a bribe, refused to pay a bribe or who met an officer who did not ask or pay a bribe. The crowdsourced reports feed into city, state and country-level databases of corruption in public services. The Indian website of I Paid A Bribe receives around 25 to 50 reports per day, with many stories of official investigations leading to suspension of officials involved. The project has now partnered with 25 other countries who have replicated the site.



sensor data

Sensor data includes information collected by physical sensors recording actions and physical changes (e.g. traffic cameras, weather sensors, ambient sensors, wearables or drones). Sensor data can provide cheap, real-time measurements of anything from pollution to crop-yield.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: ARRAY OF THINGS

Data is starting to play a role in helping cities understand and manage air quality. To try and tackle this, Chicago has begun an ambitious effort called Array of Things, to install a network of approximately 100 air quality and weather sensors around the city, and make this data available for open use for researchers and educational purposes. The sensors collect real-time data about their surroundings, including air quality, climate and noise, with added privacy controls to ensure no individual data is collected. The aim is to provide real-time detection of urban flooding, high-resolution 'block-by-block' weather and climate information, and better understand population flows to improve urban planning.



social media data

Social media data is information generated by activity on platforms such as Twitter, Facebook, Github, Meetup or other social networking sites. It can be accessed via APIs or scraped, providing real-time insights into human behaviour or social trends.



STAGE:
GATHER DATA,
INFORMATION, IDEAS

B2



CASE STUDY: PETABENCANA.ID

PetaBencana.id, is a project that combines data from hydraulic sensors with citizen reports over social media including via Twitter, to produce a flood map for cities in Indonesia. The system is programmed to react when someone in Jakarta tweets the word "banjir" (flood) and tags @PetaJkt. PetaBencana.id automatically replies, and asks them to confirm the tweet with geotagged photos. The platform then combines all incoming reports with official data from the city government to build up-to-the-minute, online flood maps, which are then made publicly available. Through producing real-time maps of urban flooding, PetaBencana.id represents a major advance on previous static PDF maps.



web-scraped data

Web-scraping is a method for extracting unstructured data from across the web, such as company websites, social media or job advertisements. Where official datasets are costly to gather and updated infrequently, web-scraping can provide more timely insights into social or economic trends.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: HEALTHMAP

HealthMap is a platform that brings together a range of disparate and informal online sources to monitor the outbreak of diseases. By scraping data from across the web, the platform is able to provide near real-time intelligence on a broad range of emerging infectious diseases. The website and app bring together data from news aggregators such as Google and Baidu News, as well social media data and validated official reports. The system then organises, integrates and visualizes potential threats. The platform was successfully used for tracking cholera in Haiti after the 2010 earthquake.



other data?

What other data sets might be particularly relevant or interesting?



open API

Application Programming Interfaces (APIs) provide the standards by which data is accessed and transferred between websites. Open APIs encourage collaboration and are built using open standards, allowing them to be freely accessed and used.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: OPENAHJO

City Councils make hundreds of decisions every month, and it can be hard for both citizens and civil servants to follow the progress of legislation making its way through government. Open Decisions is Helsinki City Council's attempt to standardise and publish all of its decision making data electronically. Data about meetings, agendas and decisions from all across the council are tagged according to the Popolo open government standard, and then made available via an open API called OpenAhjo. By publishing the data via an open API, a range of developers have been able to create digital applications that notify local residents when decisions are made on topics of interest. It has also been used by city employees to more easily search and track progress made by local politicians.



microsurvey

Microsurveys are a short, abbreviated form of surveying which typically take the respondent only a few minutes to complete. Microsurveys are often delivered by mobile phone, text message or a digital platform. Benefits include a much faster turnaround, and higher frequency of results, compared to traditional surveys.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: SAUTI ZA WANANCHI

Traditional, large-scale household surveys tend to be costly, and can take over a year to produce.

Sauti za Wananchi ('Voices of citizens') provides an alternative whereby mobile phones are used to collect information from a broad cross-section of citizens in Tanzania and Kenya.

For each survey 2,000 'panel members' are randomly selected from regions across the country, and sent mobile phones and solar-powered phone chargers to ensure that no-one is excluded. They are called and interviewed from a call centre in Nairobi once a month, with questions on a different topic each time. So far, Sauti za Wananchi has been used to gather insights into local people's living standards, perceptions of poverty and to understand people's access to safe drinking water. It claims to be Africa's first nationally representative mobile phone survey.



crowdmapping

Crowdmapping is a type of crowdsourcing which gathers data from different sources, including social media, text messages or geographic data, to provide real-time, interactive information about issues on the ground. Crowdmapping can create detailed almost real-time data in a way that a top-down, centrally curated, map may struggle to replicate.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: SYRIA TRACKER CRISIS MAP

Crowdmapping first came to international attention through its successful use in the global disaster relief movement. One example is the Syria Tracker Crisis Map, which has been used to crowdsource citizen reports on human rights violations since the beginning of the Syrian conflict in 2011. The map attempts to provide more detailed metrics on fatalities, while preserving the name, location and details of each victim. The service blends reports from local news with on-the-ground reports, using hashtags on social media or sent via email. Nearly 5,000 submissions, including over 11,000 fatalities, have been reported since the map's launch, with collected data being used by both USAID and the Washington Post to report on local events.



citizen science

Citizen science is any process where scientists and (usually unpaid) volunteers work together to collect or process scientific data or observations. Citizen science unlocks new resources for research, experimentation and analysis by opening the process to everyone.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: MOSQUITO ALERT

The World Health Organisation reports over 500 million cases of mosquito borne illnesses a year, but the global spread of disease is difficult to manage or track in real-time. The Global Mosquito Alert Consortium (GMAC) was established in 2017 as a global network of citizen science projects that follow a common set of four protocols to track breeding sites, bites and sightings of mosquito species known to carry diseases. One such initiative includes Mosquito Alert, a project based in Spain that works with local communities and schools to build capacity for citizen science. It provides an app for volunteers to submit pictures of mosquitos or local breeding sites. These are then validated by a team of experts, and presented on an interactive map. Since 2014, the platform has been downloaded 57,000 times, reporting over 12,000 observations of tiger mosquitoes and 3,117 breeding sites.



data collaborative

Data collaboratives are a form of collaboration in which partners from different sectors - including private companies, research institutions, and government agencies - enter into an agreement to exchange data for a specific social cause. Data collaboratives aim to unlock new value in private data that would not be exploited otherwise.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: UN GLOBAL PULSE

In recent years the Mexican state of Tabasco has experienced record-breaking rainfall, yet timely data about how floods are affecting the population in real-time is currently missing. In response, UN Global Pulse brought together a range of novel datasets to generate real-time insights about human behaviour during flooding events. Data sharing and analysis was made possible via a public-private partnership between Telefonica Research and data scientists from the Technical University of Madrid, under guidance from experts at Global Pulse and the World Food Programme. By combining citizen mobile phone data, remote sensing data (satellite images), rainfall data, and census data from citizens, the project was able to use patterns of mobile phone activity to highlight the impacts of flooding on infrastructure and the local population, whilst also helping to target and improve public communications for disaster response.



solution mapping

Solution mapping is a method that helps organisations and the communities they work with to identify needs, issues and opportunities by looking for solutions developed by people in response to concrete problems they face.



STAGE:
GATHER DATA,
INFORMATION, IDEAS



CASE STUDY: HONEY BEE NETWORK

In many regions across the world there is tremendous potential for development through building upon what people already know and their capacity to learn from one another. The Honey Bee Network, based in India, aims to pool local solutions and facilitate communication among farmers, artisans, pastoralists and other grassroots innovators all over the world.

One of the network's methods for gathering solutions involves 'innovation scouting', where university students are asked to survey local villages for local innovations during their summer holidays. The network has built what is now one of the largest databases in the world on farmers' innovations. It is designed to be easy for local communities to use, and provides multiple translations in local languages. A key principle of the platform is that knowledge holders must benefit from the success of their innovations, in both fame and remuneration.



policy makers

Government and regulators



STAGE:
MOBILISE PEOPLE



entrepreneurs

Individuals or small and medium
sized enterprises



STAGE:
MOBILISE PEOPLE



central government officials

Civil servants and officials



STAGE:
MOBILISE PEOPLE



non government organisations

Charities and not-for-profit organisations



STAGE:
MOBILISE PEOPLE



local community

People within the neighbourhood



STAGE:
MOBILISE PEOPLE



investors

Funders or shareholders



STAGE:
MOBILISE PEOPLE



employees

People who work directly with or for you



STAGE:
MOBILISE PEOPLE



professionals

Doctors, nurses, lawyers, teachers, police,
armed forces etc.



STAGE:
MOBILISE PEOPLE

C2



retired people

Older people who may have retired



STAGE:
MOBILISE PEOPLE



other?

Who else might be particularly relevant
or interesting?



STAGE:
MOBILISE PEOPLE

C2



deliberation

Deliberation is a method of weighing up different options through dialogue. It can typically be used for engaging members of the public or participants in considering trade-offs, and in establishing greater understanding of one another's views to arrive at more consensus-driven and informed set of solutions or recommendations.



STAGE:
CONNECT & INTERPRET



CASE STUDY: ESTONIAN PEOPLE'S ASSEMBLY

In 2012 the ruling Reform Party in Estonia was beset by a scandal related to anonymous donations, which sparked a national engagement exercise. The five-stage process began with crowdsourcing of over 2,000 proposals on an online platform, on topics from reducing barriers to creation of political parties, to improving party conduct. The process then culminated in a face-to-face deliberation between randomly selected citizens. Over several days, the group was tasked with sorting through the online proposals, and making practical recommendations to the Parliament. Fifteen proposals were submitted, leading to three new items of legislation and another four proposals being partially adopted.



collaborative platform

A collaborative platform enables a loosely defined group of participants to come together to create a product or service, which is typically then made available to contributors or non-contributors alike. It draws on principles of open collaboration and open source software development, whereby openness and transparency in product development improve participation, trust and collective memory.



STAGE:
CONNECT & INTERPRET



CASE STUDY: WIKIFACTORY

Wikifactory is an open platform that allows product designers and creators to collaboratively solve problems using digital fabrication technologies. It responds to a problem among product designers where sharing PDFs and different file types over email, or via a shared drive, hinders effective collaboration. The platform provides something similar to Github for product designers, including version-control, an issue and documentation tracker to crowdsource ideas or solutions, a permission system and community tools to aid collaboration. By opening up the process of product design, Wikifactory also makes it possible to invite collaboration on designs from creators all over the world, who may copy, use or replicate the designs. Many projects currently on the site aim to provide cheap solutions for solving social problems using digital fabrication, from small wind turbines to water filtration systems for clean water. So far over 3,000 product developers from around the world have signed up to the platform.



open source repository

An open source repository is a digital repository where content (e.g. code, text or other learning resources) can be stored and freely downloaded with few restrictions on use. Many open source repositories help aid collaboration by providing a space for uploading documentation, monitoring and version control.



STAGE:
CONNECT & INTERPRET



D1

CASE STUDY: GITHUB

Software projects can be complex and often require contributions among a wide variety of developers and teams. Github is a platform for enabling collaborative, non-linear working with open source files. The platform is built on top of the 'Git' distributed version control system whereby every user contributing to a project holds a version of the project's content (a 'repository') which they can edit, add to and test. When they are satisfied with their edits they can 'commit' them, and the system exchanges sets of changes, synchronising everyone's latest work.

GitHub hosts these repositories and adds a user interface and tools that help with coordinating work. The platform makes detailed statistics of users' work open to inspection by default. Github reports having over 30 millions users and 100 million repositories, making it the largest host of source code in the world.



data warehouse

A data warehouse is a central database optimised to analyse information coming from a range of different sources. Data is cleaned, organised and structured in advance, acting as a single source for querying and enabling faster decisions.



STAGE:
CONNECT & INTERPRET



CASE STUDY: THE NEW YORK CITY FIRE DEPARTMENT

The New York City Fire Department (FDNY) uses a Risk-Based Inspection System in order to better understand buildings that are most vulnerable to fire across the city. Collecting data from building inspections was previously done by paper, making it prone to delays and difficult to share information across the city's various fire companies, battalions and divisions. Therefore the department has built a data warehouse that all internal partners can access. This has helped to streamline the FDNY's entire inspection workflow, allowing statistics from the inspections to be gathered in one place, and improving measurement against internal goals.



online forums

Online forums are platforms that allow people to share ideas and resources, to deliberate on key issues or simply to talk to one another. They may be used by policymakers to generate discussion or suggestions on a bill or document, or they may be set up by projects or companies trying to build and sustain a community.

Most successful forums involve active and ongoing moderation.



STAGE:
CONNECT & INTERPRET



CASE STUDY: SMARTPARTICIPATION

SmartParticipation is an open source e-participation platform designed to offer an adaptable online forum environment for informed discussion and ideation.

The design was informed by experiments run by researchers at the University of Cornell in partnership with several government agencies in the US. The platform itself provides easy content authoring, targeted commenting and options for moderating discussions.

One of the many experiments run on the platform was a 2011 consultation on Airline Passenger Rights in partnership with the Department of Transportation (DOT). More than 19,000 visitors posted over 900 comments on a proposal that would give airline passengers increased protections. After the comments were summarised and sent to the department, DOT announced that they played an important part in shaping the final rules.



study circle

Study circles are groups that meet regularly with an aim to develop collective social expertise, and to brainstorm solutions or ideas. They try to incorporate diverse perspectives and encourage people to respectfully consider the views and values of other members within a community.



STAGE:
CONNECT & INTERPRET



CASE STUDY: RESILIENCE DIALOGUES

The Resilience Dialogues were a set of facilitated discussions in the US designed to help communities create climate action plans. The service uses an online platform, developed by the Massachusetts' Institute for Technology's Climate CoLab. Ideas from communities are matched with expert knowhow, and online discussions are used to clarify local risks and opportunities, share strategies that have worked in other communities, and build shared commitment to future plans and actions.

Local knowledge, combined with prompts from experts about the resources needed for various ideas and the likely outcome of particular proposals lead to much more robust proposals than either of these groups could have produced in isolation. Communities leave the dialogues with more confidence, as well as a range of new resources and a strategy for implementing resilience efforts in their local context.



hackathon

A hackathon is a challenge-based event where the organiser defines a problem or challenge, and participants compete to create a solution. Sometimes this involves a reward. Hackathons bring people together to help surface new ideas or solve a problem quicker.



STAGE:
CONNECT & INTERPRET



D1

CASE STUDY: TAIWAN PRESIDENTIAL HACKATHON

In 2018, the Taiwanese government launched the Taiwan Presidential Hackathon to demonstrate its commitment to open source and open data, and to accelerate solutions that address the needs of the country by inviting social innovators to propose project ideas using data and technology.

The government provided participating teams with open government datasets, as well as some 'protected' data that can be made accessible upon request. Submissions were judged based on criteria: innovation, social influence and feasibility. In 2019 ten finalists were selected from over 100 submissions, with solutions ranging from predictive monitoring of water leakages to a platform for improved knowledge sharing among caregivers. The government has now introduced an international track that invites submissions from all over the world.



proxy indicator analysis

In statistics, a proxy is a variable that is not directly relevant, but can serve in place of an unobservable or measurable variable. Proxy indicator analysis is useful for inferring a measurement for unavailable data based on known correlated information.



STAGE:
CONNECT & INTERPRET



D5

CASE STUDY: UNDP SUDAN

The UNDP office in Sudan tested the use of satellite data for estimating poverty levels via changing night-time energy consumption. The team used data pulled from night-time satellite imagery and analysed these illumination values over a two-year period, in conjunction with electric power consumption data provided by the national electricity authority. The proof-of-concept showed that night-time satellite imagery provided a reasonable proxy for poverty, though more detailed analysis with higher spatial and temporal resolution would have provided a better result. The experiment demonstrated the feasibility of using satellite imagery and machine learning to understand poverty, which a number of other organisations have continued to explore.

predictive analytics

Predictive analytics encompasses a variety of statistical techniques that analyse current and historical facts to make predictions about future or otherwise unknown events. It can be used to identify early warning signs or to predict likely outcomes based on a range of variables.



STAGE:
CONNECT & INTERPRET



D5

CASE STUDY: AIME

Artificial Intelligence in Medical Epidemiology (AIME) is an initiative that uses data from a variety of different sources to predict disease outbreaks, providing warnings before the spread of disease occurs. The system relies on an algorithm that blends over 20 different variables from local government data and satellite data, with variables such as weather, wind speed and direction to rainfall. The result of this analysis is a map which illustrates the probability and location of the next outbreak. Following initial tests in Malaysia, Brazil and other South American countries, AIME's successes include an algorithm that can predict the location of the next Dengue, Zika or Chikungunya outbreak up to three months in advance with over 86 per cent accuracy.

crowd forecasting

Crowd forecasting is a method that asks small or large groups to make predictions about the future. It is inspired by research which showed that small crowds of non-experts can often forecast political events more successfully than individual experts.



STAGE:
CONNECT & INTERPRET



CASE STUDY: EARLY WARNING PROJECT

Genocide and mass atrocities are almost always preceded by a range of early warning signs. The Early Warning Project aims to improve the early warning system for mass atrocities using a mixture of crowd forecasting and statistical modelling. An annual comparison survey invites experts in the field to participate, presenting them with pairs of countries and then asking them to choose which is more likely to experience a new mass killing. The cases from the annual survey then inform the selection of countries which the Early Warning Project tracks in real time using a crowd prediction polling tool called Good Judgment Open. Anyone can sign up to take part and the insights from the crowd consensus help address gaps in the statistical model. Early Warning Project produces a ranked list of more than 160 countries based on their likelihood of experiencing an onset of mass killing in order to better target preventative action.

heatmap

A heatmap is a graphical representation of data where values are represented as shades or colours, overlaid on a matrix or map. It can help to quickly make sense of relationships between different variables, or trends presented geographically.

CASE STUDY: CEASEFIRE IRAQ

Ceasefire has a civilian monitoring tool to monitor violations of human rights or international humanitarian law. People can submit reports of violations occurring anywhere in Iraq. Submitted reports are stripped of any personal identifying information and plotted onto a live map showing the distribution of violations by location and type. It is used to create a more accurate and up-to-date picture of the human rights violations on the ground, improving accountability and motivating a more effective national and international response.



STAGE:
CREATE CHANGE



what works centres

What Works Centres aim to ensure that public spending and practice is informed by the best available evidence. Centres operate in their own distinct way, but all aim to generate evidence and translate it into relevant and actionable guidance.

CASE STUDY: THE EDUCATION ENDOWMENT FOUNDATION (EEF)

The EEF is an independent charity dedicated to breaking the link between family income and educational achievement, by gathering research evidence and making it actionable to teachers and practitioners across the UK. EEF's Teaching Learning

Toolkit is an accessible and interactive web tool that summarizes research evidence for teachers and schools on how to improve the attainment of disadvantaged people, according to cost, evidence and impact. More than half of secondary school leaders in the UK now say they use this resource.



STAGE:
CREATE CHANGE



E1

emergency alert

Emergency alerts improve awareness or speed up the response to an emergency or event. They often involve blending or integration of data, analysis and some form of prediction of the likelihood that an event will happen. This information is then presented on a map, sent via SMS, radio or email.



STAGE:
CREATE CHANGE



CASE STUDY: VAMPIRE

The combination of El Niño - a severe and lengthy climatic event exacerbated by climate change - and severe drought in districts across Indonesia gravely intensified threats to food security in Indonesia. The

Vulnerability Analysis Monitoring Platform for the Impact of Regional Events (VAMPIRE) was designed by Pulse Lab Jakarta and the World Food Programme to provide integrated map-based visualisations of how drought will affect local populations, bringing together household survey data, vegetation health, rainfall data and other indicators. Based on the measure of economic vulnerability and exposure to drought, the tool can alert authorities to areas where people may require assistance. It has been embedded into the situation room of the Office of the President of the Republic of Indonesia. Further improvements may include the integration of text-messaging alerts to notify farmers or other affected populations.

register

Registers are lists or datasets that act as core reference data, providing lists of things like schools, companies or job centres. Registers can help to coordinate tasks across diverse crowds, in turn improving collective memory and making collaboration easier across time and space.



STAGE:
CREATE CHANGE



CASE STUDY: EPISTEMONIKOS

Epistemonikos is allegedly the most comprehensive database of systematic reviews and other types of scientific evidence about health conditions and treatments. Initially, machine learning algorithms search through the major global health related databases to generate a longlist of potentially relevant articles, and assign scores to indicate certainty based on their relevance to the question. This longlist is then analysed for inclusion by at least 2 independent 'expert' humans in the loop, using a Tinder-like app. A validated dataset is then labelled and retained on the online platform for future use by the clinical community. The Chilean Ministry of Health issued 22 health guidelines within 6 months following an evidence review using the Epistemonikos platform.

