

## **data cards**



## 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).



### 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.

## official data

Official data can include longitudinal survey data, such as census data, electoral registers or other population data. Official data provide a long-term picture of a country over time, and are usually collected by public institutions or national statistical agencies as a public good.



### CASE STUDY: AFRICAPOLIS

Africapolis aims to be the most comprehensive and standardised geospatial database on cities and urbanisation dynamics in Africa. It blends data from across the continent in an effort to provide standardised, comparable insights into Africa's fast-changing urban populations.

The single most important element is official population records, including census data of official cartographic resources such as administrative boundaries. In some cases official data can date back 30 or more years, which is significant given the pace of demographic and urban trends. Africapolis brings these datasets together into a single, comparable resource, as well as additional sources such as satellite data to confirm the precise location of settlements. By integrating thousands of smaller agglomerations (covering inhabitants of a total of 180 million people) Africapolis also highlights the role of smaller towns and cities in urban networks, the emergence of new urban areas and the increasing fluidity between urban and rural environments.

## 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.



## satellite data

Satellite data is aerial imagery collected by satellites, usually available from companies for a fee. It can provide detailed insights into land-use over time, or act as a proxy for other measures.



### 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.

### CASE STUDY: REGEN NETWORK

Regen Network collects ecological and agricultural data via remote sensing (images collected by satellites, drones, and planes), and on-the-ground observation via Internet of Things (IoT) sensors. The collected data is verified by farmers and then run through a trained machine learning model which tries to determine ecosystem health. So far, Regen Network have developed an algorithm that is capable of detecting till events in Romania and Ukraine with a 99 per cent accuracy through public satellite data.

## 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.



## open data

Open data is the raw data that is gathered by people or organisations, published in an electronic format that machines can read; it is then shared online and allowed to be re-used by others instead of keeping it private.



### 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.

### CASE STUDY: OPENCORPORATES

OpenCorporates increases transparency in the corporate world through making information about companies more accessible, so that citizens and journalists can better monitor and regulate them. OpenCorporates crowdsources data from citizens, who contribute to populating and updating the platform, identifying errors, or importing web scraped data. The contributions from people all over the world and their local knowledge have made OpenCorporates the largest open database of company data in the world, with over 100 million companies in a large number of jurisdictions. The fact that the data is open has also improved data quality, allowing anomalies, errors and issues to be discovered faster.

## 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



## ethnographic data

Ethnographic data encompasses qualitative information (stories, personal accounts or transcripts) collected by interviews, participant observation, note-taking, online text or recording. It provides rich and highly detailed accounts of people's lives.



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.

### CASE STUDY: ON OUR RADAR

On Our Radar trained citizen journalists to provide detailed accounts of the Ebola crisis from communities all around Sierra Leone (e.g. from urban slums, polio camps, and remote rural villages). The reporters shared their reports with On Our Radar via recorded telephone interviews, WhatsApp audio and photo functions, and SMS. People in the poorest regions of Sierra Leone are more likely to have a mobile phone than access to basic amenities and electricity, so even those in the most remote regions could report.

On Our Radar reports captured the atmosphere of day-to-day life in vulnerable communities in Sierra Leone, often delivering breaking news that international media did not otherwise have access to, and from communities that Western journalists failed to reach. Its reports were featured on the BBC World Service, Guardian and Sky News among others.

# gather data method cards

INSTRUCTIONS:



## crowdsourcing

Crowdsourcing is an umbrella term for a variety of approaches that source data, information, opinions or ideas from large crowds of people, often by issuing open calls for contribution. It can help bring new ideas to light that hadn't previously been considered, or to gather expertise from people who have specialised knowledge or understanding of an issue.



## 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.



### CASE STUDY: DECIDE MADRID

The City Council of Madrid has used CONSUL, an open software tool, to create a citizen platform - Decide Madrid - to crowdsource citizen proposals and harness more local knowledge in decisions about how to plan and allocate budgets. New ideas from residents have been gathered and implemented on topics ranging from sustainability and air pollution, to ticketing for local transport. Since its creation it has allocated over 100 million Euros in funding for projects suggested by local people. On the platform, people can make proposals, engage in debates, and vote for ideas. It is often used in conjunction with offline engagement.

### 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.

## challenge prize

Challenge prizes or competition platforms are a method of drastically widening the pool of possible solutions to solve a problem. They are useful for getting a variety of responses to questions and of spurring novel suggestions for how to address an issue or specific challenge.



## 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.



### CASE STUDY: DARPA GRAND CHALLENGE

In the US, the Defense Advanced Research Projects Agency (DARPA) ran a grand challenge with the goal of supporting innovation to accelerate the development of autonomous vehicle technologies with military applications. No vehicles completed the set route for the \$1 million prize in 2004, but a year later another round was held in the desert near the California/Nevada state line. A total of five teams completed the course which was 132 miles over the desert terrain. The Stanford Racing team took home the \$2 million prize, completing the course in 6 hours 53 minutes. DARPA credits this challenge with the creation of 'a mindset and research community that a decade later would render fleets of autonomous cars and other ground vehicles a near certainty for the first quarter of the 21st century'.

### 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.

## gamification (serious games)

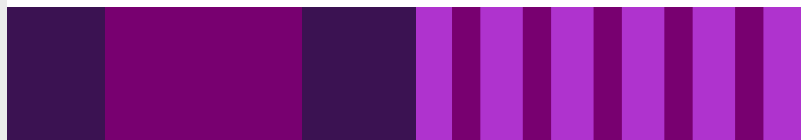
Gamification is a broad term that refers to using game-like elements to make engagement in collective intelligence projects more fun. It can be a useful way to motivate audiences in complex topics or research, as well as illustrating trade-offs associated with making certain choices. Games can also be useful for presenting data, ideas and trends.



STAGE: B3  
GATHER DATA,  
INFORMATION, IDEAS

### CASE STUDY: SEA HERO QUEST

Sea Hero Quest is a mobile game where players' actions help scientists to understand and fight dementia. The game asks users to find their way through a digital maze, in turn providing researchers with valuable data to understand the user's spatial navigation. According to its website, playing Sea Hero Quest for only two minutes generates the same amount of data that it would take 5 hours to collect in similar lab based research. As a result the project has gathered approximately 17,600 years worth of dementia research from players on the app (from around 4.3 million players). Initial results from the data have provided novel insights into spatial navigation abilities across different social groups, ages and countries.



## crowdmapping

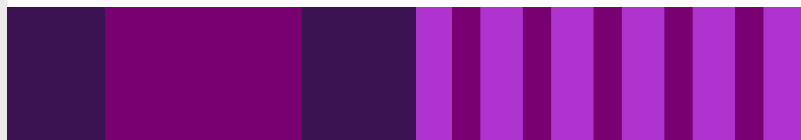
Crowdmapping is a type of crowdsourcing which gathers data from different sources, including social media, text messages or geographic data, to provide real-time, interactive information about issues on the ground. Crowdmapping can create detailed almost real-time data in a way that a top-down, centrally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

### 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.



STAGE: B3  
GATHER DATA,  
INFORMATION, IDEAS



## participatory sensing

Participatory sensing is where citizens use lightweight, cheap sensors to collectively monitor the environment around them. It also includes the task of collectively sharing and interpreting streams of citizen-sensed data with other community members, deepening their understanding of the issue, educating participants and empowering them to act.



**CASE STUDY: IDEAS FOR CHANGE**  
Noise pollution is a problem which affects many people living in inner city Barcelona. A think tank called Ideas for Change worked with local communities to help them install sensors that measure the harmful effects of noise pollution in their neighbourhood. Participants were provided with cheap, open-hardware sensors, and then guided through the process of setting-up the technology and sharing data with one another.

The council organised public meetings in which locals could talk through their findings and propose potential solutions. Some of the solutions that have been implemented include new flower beds that remove areas where people used to sit and drink into the late hours of the evening, and improved community policing.



## 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.

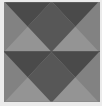


**CASE STUDY: MOSQUITO ALERT**  
The World Health Organisation reports over 500 million cases of mosquito borne illnesses a year, but the global spread of disease is difficult to manage or track in real-time. The Global Mosquito Alert Consortium (GMAC) was established in 2017 as a global network of citizen 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.



**incentives  
and retention**





## incentives and retention

What might motivate and incentivise people to be involved?

| ENGAGE                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                                                |                                                                                                                                                                                                                                                                                       |                                                                                                                                                                                                                                                                 | RETAIN                                                                                                                                                                                                                                                                     |  |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|
| <b>FINANCIAL REWARD</b><br><br><b>Winner takes all prize</b><br>People are motivated to contribute by the opportunity to gain income for themselves or their team.<br>Eg. Challenge Prize | <b>COMPENSATION</b><br><br><b>Task compensation</b><br>Individuals are motivated by the guarantee of cash or other rewards (such as vouchers) for completion of a task.<br>Eg. Amazon Mechanical Turk       | <b>REPUTATION</b><br><br><b>Gaining recognition</b><br>Individuals are motivated by the thought of having their knowledge and expertise recognized by their peers.<br>E.g. Crowdsourcing platforms | <b>FEEDBACK</b><br><br>Good feedback is immediate, interpretable and identifies the usefulness of contributions made by participants. This can help them develop their skills. | <b>COMMUNICATION</b><br><br>Good communication is up to date, accessible in the participants' own time, two way, and doesn't bombard people.                                              |  |
| <b>PURPOSE</b><br><br><b>Greater purpose</b><br>People are motivated by the prospect of helping to move an idea or area of research forward in ways they couldn't alone.                  | <b>PASSION</b><br><br><b>Fulfillment of passion</b><br>People are motivated to work on areas of interest which align with their passion or values.                                                          | <b>SATISFACTION</b><br><br><b>Satisfaction</b><br>Individuals are motivated by the prospect of being able to do something well, commensurate with their skills.                                    | <b>GUIDANCE</b><br><br>Good guidance ensures that contributors understand where and how they are best able to contribute. This can include FAQs or clear instructions.         | <b>BENEFITS TO CONTRIBUTOR</b><br><br>Make explicit and prioritise what the benefits are for contributors, not only the project and the stakeholders/beneficiaries.                       |  |
| <b>LEARNING</b><br><br><b>Personal development</b><br>People are motivated by the idea of improving themselves; developing a new skill or gaining new knowledge.                        | <b>SOCIAL CONNECTIONS</b><br><br><b>Growing a community</b><br>Individuals are motivated by the opportunity to build or expand their connections with likeminded people.<br>E.g. Peer to peer communities | <b>OTHER</b><br><br><b>???</b><br>What else might incentivise and motivate people to participate?                                                                                                | <b>ORGANISATION</b><br><br>Demonstrating high levels of organisation will give your contributors confidence in the process and in the commitment they are making.            | <b>INTEREST IN PARTICIPANTS</b><br><br>Demonstrating genuine interest in your contributors and the value they provide will reassure them they are not being used or wasting their time. |  |

