



OPEN FOR
FEEDBACK



WORK IN
PROGRESS

Exploring the Covid-19 symptom tracking ecosystem

Contents

Contents	1
About	2
Executive summary	3
Introduction	5
Landscape review	5
Key findings at this stage	6
A number of symptom trackers in the ecosystem	6
Ecosystem structure	6
Lack of collaboration	6
Lack of transparency around data collection	6
Open data availability	6
User base	7
Specific types of users	7
Location	7
Data discovery	8
Types of data	8
Observations about types of data	9
Commonly asked questions	9
Data standards	9
Representation of communities	10
Changes over time	10
Mental health and wellbeing	11
Outputs	11
Individual guidance	11
Research insights	11
Improvements to existing services	11
How does the data collected relate to the outputs?	12
The symptom tracking data ecosystem	13
The UK symptom tracking data ecosystem	13
Where is value being created in this example?	14
Barriers to impact	14
Opportunities	15
International variations	15
A single source approach	15
A more collaborative approach	16
Where is value being created in this example	17

Discussion	18
Data discovery	18
Legal complexities	18
Technical barriers	20
Trust and trustworthiness	20
Recommendations	20
Standards for data	21
Enabling hub to facilitate interactions	21
Building trust with the public through ethical approaches	22
Appendices	23

About

This report has been researched and produced by the Open Data Institute, and published in January 2021. Its lead author was James Maddison with contributions from Olivier Thereaux and Jeni Tennison. If you want to share feedback by email or would like to get in touch, contact the Open Covid-19 project team at research@theodi.org.



How can it be improved? We welcome suggestions from the community in the comments.

Executive summary

In April 2020, the Open Data Institute (ODI) [launched a new project](#), funded by Luminate, which aims to support the building of an open and trustworthy data ecosystem for Covid-19. As part of this project, our team looked at the ecosystem of symptom trackers to understand how they operate and cooperate, and what data flows they enabled.

Covid-19 symptom trackers may be expected to collect similar data and have a comparable purpose, but our team observed a diverse ecosystem of organisations, working towards a variety of outcomes.

We did find instances of organic collaboration between symptom trackers, but coordination seemed to be rare. It was generally limited to government intervention – either by mandating a single tracker, or by empowering an organisation with an intermediary role to manage the aggregation and processing of data

While this disaggregated, decentralised approach was generally more agile, and allowed the emergence of symptom-tracking applications at a time of crisis and emergency, there would have been value in interventions aimed at increasing standardisation, interoperability, and knowledge sharing.

This document is not a comprehensive review of the symptom-tracking data ecosystem. It provides an overview of our research in this area and outlines some high-level recommendations for symptom trackers, health policy organisations and funders.

Introduction

During the early stages of this project, the ODI [offered free support](#) to people and organisations who have data they think could help the world to navigate the pandemic. One organisation which requested support was the team behind TrackTogether, an application which allows users to monitor and record symptoms of illness.

These types of applications – known as [symptom trackers](#) – can help individuals to decide whether they have symptoms of Covid-19, and therefore need to seek further medical advice or support. If enough users engage with a particular symptom tracker, the data that is being collected could be used by researchers for a number of purposes, for example to identify where Covid-19 outbreaks are likely to occur next. These types of insights could be used by decision makers across sectors to make more informed decisions about how, when and where to respond to the pandemic. Symptom trackers are not a new phenomenon, but have become an integral part of how we learn about the spread of Covid-19 and its symptoms.

The ODI helped [TrackTogether](#) to publish open data from their symptom tracker, which involved using:

- the ODI's ['Anonymising data in times of crisis' guide](#) to reduce the risks of reidentification of individuals through the data and
- the [Octopub](#) open data publishing tool to make the data available on Github. Aggregate TrackTogether data is [now available](#) for anyone to access, use and share.

The work with TrackTogether highlighted a number of interesting questions:

- How many other symptom trackers have been created to help navigate the pandemic?
- Are these applications collecting similar data from users, and what is this data being used for?
- Are the organisations behind these applications working together to help address the wider social and economic challenges presented by Covid-19?

With these questions in mind, the ODI decided to [make exploring the symptom-tracking data ecosystem a key area of focus](#) for the Covid-19 project.

The research was broken down into three phases:

1. Landscape review of existing symptom trackers
2. Data discovery
3. Data ecosystem mapping

Landscape review

Landscape reviews provide a broad description and analysis of policy areas, largely based on publicly available information.

In phase one of our research, we conducted a landscape review of the Covid-19 symptom trackers that are currently in operation. Initially, we found 38 applications that are collecting and using data about user's symptoms. We used targeted web search terms such as 'Covid-19 symptom-tracking' or 'Covid-19 symptom trackers', as well as questions such as 'have I got Covid-19?'. Our colleagues at TrackTogether helped by sharing a list of organisations that they were aware of who collect Covid-19 symptom data.

To understand whether these applications met with the team's definition of a symptom tracker – an application that helps you monitor and track the symptoms of Covid-19 – each initiative was described in terms of the following attributes:

- **Initiative name**
- **Lead organisation**
- **Lead organisation type** We defined four broad types of organisation: private; public; third [sector] and research. It felt important to note this so that we could examine an organisation's motives later, following data discovery.
- **Connected organisations** Information about other organisations involved in an initiative. This could include technical partners, users of the data, those with an interest in the outputs being created, and funders. This helps understand how organisations are sharing data and whether there are any common denominators across different initiatives.
- **Country:** This helped us track where the user base for each application was, which is useful for understanding the ecosystem that surrounds different initiatives.
- **What data is being collected?** This attribute captured any public information about data being collected by the application. This helped us identify where to start the data discovery process.
- **Is data being shared with others?** Any publicly available information about data being shared as part of an initiative. It felt important to try to understand if applications were communicating to users about data sharing, as well as discovering whether any initiatives were sharing with each other.
- **Outputs** Any information about publicly-accessible outputs relating to Covid-19 that have been created by the organisations involved in a particular initiative.
- **Website URL**

The full dataset of symptom-tracking applications that we identified is available in a [spreadsheet](#).

Key findings at this stage

A number of symptom trackers in the ecosystem

This research does not capture a complete list of all symptom trackers in operation around the world, but it does highlight that there are a large number of symptom trackers for users to choose from, particularly in the UK and the US.

Ecosystem structure

In most countries, including the UK, it does not appear that there is just one accredited symptom tracker which has been endorsed by government or national healthcare systems. There are some exceptions to the rule: in Canada and Vietnam, governments seem to favour a more centralised approach and publicly endorse one symptom-tracking application.

Lack of collaboration

We found only two large scale initiatives – [Project Oasis](#) and the [Coronavirus Census Collective](#) – that had the same organisations formally involved (beyond the NHS as a user of the insights being created). This suggests a lack of collaboration between symptom trackers.

There are contexts in which a very diverse and independent ecosystem is ideal, but we expected to see more collaboration in the Covid-19 ecosystem, given the need to respond to the pandemic quickly and efficiently. In this context, having numerous data stewards who do not collaborate may lead to fragmentation, leading to bad decisions that affect people's health and wellbeing.

Lack of transparency around data collection

From the outset, we did not find any public information about the types of data that different initiatives were collecting. The large majority of initiatives have published little or no information about the datasets that they are collecting. While this might come as a result of the need to deliver something quickly and efficiently, transparency of activities around public health data is important for public trust, and should not be forgotten or deprioritised.

Open data availability

Beyond TrackTogether's open data, we found only two initiatives publishing open symptom-tracking data. They were [symptomtracker.org.uk](#), a now defunct volunteer-led symptom tracker, and [Studybugs](#), an application for parents to notify schools when their child has Covid-19 symptoms.

User base

16 of the symptom trackers can be accessed using a web browser and 12 symptom trackers are available as a downloadable application on a smartphone. 10 of the symptom trackers have been integrated into existing health and wellbeing applications, which consumers use to track personal information on subjects such as fitness, diet and even medical records. According to [a study](#) published in 2019, around 33% of the global population over the age of 15 use health tracking applications or devices, a large majority of which are between the ages of 20 and 39. It is therefore reasonable to surmise that the majority of symptom trackers are targeted at people within a similar age range, that have access to smartphones or a computer.

Specific types of users

Most of the symptom-tracking applications are marketed for a broad audience, but some have been created for very specific audiences. Studybugs, for example, is designed for use by parents of school aged children.

Location

The majority of applications examined are marketed as tools that can be used by citizens within a given geography, for example people in the UK. Some applications require proof of residency before you can use them. Two of the initiatives – one in Ghana and one in the US – require you to submit a country-specific phone number. Failure to do so means a user cannot use the application. Three of the initiatives examined have publicly stated that they are operating on a global scale, so users from multiple countries are welcome to use the application.

Data discovery

In order to identify the types of data that users are being asked to share with symptom-tracking applications, we accessed each application and filled in the questionnaires provided. At this stage, the team was very careful to not submit any false information to the questionnaires, so this research might not reflect all the questions asked by some applications if they operate using a question tree whereby questions are revealed based on answers to previous questions.

It is also worth noting that at this stage, 14 of the initiatives were ruled out for data discovery (see appendix). Two of these initiatives were re-using symptom-tracking data from other sources to produce insights on topics such as mobility during the pandemic, but were not directly tracking user's symptoms, and for a number of the non-UK applications, the team were unable to access the questionnaires due to a requirement to provide proof of residency, such as a phone number or postcode. However, we did not rule out these initiatives for the data ecosystem mapping exercise below, as some of them are integral to the ecosystem.

Types of data

Based on the 24 symptom trackers that we accessed, symptom-tracking data falls into 11 broad categories:

- Personal and demographic
- Employment
- Household
- Location
- Pre-existing health
- Symptoms of illness
- Covid-19 infection status
- Travel
- Interaction with others
- Changes in lockdown
- Mental health and mood

The full list of questions is available in an aggregated list in the appendices of this report. There is also a breakdown of how many initiatives asked each question.

Observations about types of data

Commonly asked questions

The five most commonly-asked questions require the following information:

- The user's symptoms that correlate with symptoms officially associated with Covid-19 (16/28)
- The age of the user (16/28)
- Exposure to someone who has been diagnosed with Covid-19 (13/28)
- The user's postcode (12/28)
- The user's pre-existing health conditions, such as heart disease, diabetes, asthma, etc (12/28)

Data standards

Many of the applications are collecting the same types of data, but there is a lot of variation in how different applications ask users to provide information. The most prominent example of this in our research is the way that symptom trackers are documenting the age of users. A user's age is one of the most commonly asked questions, but some trackers ask for a full date of birth, others ask for a specific age and many ask which age category the user falls into. This suggests that there are few commonly adopted data standards – documented agreements about how to represent data, which enable us to take a consistent, rule-based approach to data management – for Covid-19 symptom data.

In order to understand data, to share it with others and to help them understand it too, the data needs to conform to a widely recognised format. Health organisations like the National Health Service (NHS) recognise the importance of data standards, and are [working towards](#) making sure that information about health is captured and categorised using common standards.

Representation of communities

Given that Covid-19 affects different groups and communities in different ways, it is important to make sure that the data being collected is representative of all of the communities in the area that the application is serving. Gathering data from members of groups that have been highlighted as being at a higher risk of fatality from Covid-19 is particularly important as an early warning signal.

Early examinations of Covid-19 cases and deaths in multiple countries have found that men are [more at risk of severe illness and death](#) than women, so it's useful that eight of the symptom trackers examined are collecting data about an individual's assigned sex. However, four of the symptom trackers also ask users to disclose their gender, and one asks for users to disclose their gender without asking questions about assigned sex. It is not clear whether questions about gender are being asked to address a specific research question, or because gender and sex have been amalgamated into one attribute.

One very evident omission by many symptom trackers is questions around race and ethnicity. Only four of the identified trackers asked at least one question about ethnicity and only one asked multiple questions. Research shows that Covid-19 can have a different impact on people from different communities. For example, [early research into ethnicity and Covid-19](#) suggests that in the UK, the Bangladeshi population has been disproportionately affected by Covid-19, with death rates nearly twice as high as the White British population.

Given this evidence that the effects of Covid-19 vary for people from different communities, more symptom trackers should consider asking questions about [protected characteristics](#), as well as socioeconomic status so that the valuable insights created by symptom trackers do not exclude people from different communities. For more information about why it is important to collect information about protected characteristics for services that benefit the public, see our report '[Monitoring Equality in Digital Public Services](#)'.

Changes over time

10 of the symptom trackers that we looked at ask questions about how lockdown (as opposed to Covid-19 itself) has affected the user, and whether they have seen noticeable changes over time. This includes areas such as diet, weight, physical activity and sleep.

Five of the initiatives asking these questions are private-sector-led, which might suggest that symptom trackers are also being used as an opportunity for companies to collect data to help inform their product or service offerings in future. The trackers in question offer no clear picture of how this data will be used, and who it will benefit. Given that many symptom trackers are collecting data by leveraging public goodwill, they should strive to be as transparent as possible about what this data is being used for, in order to avoid any breaches of public trust and comply with existing data protection laws.

Mental health and wellbeing

11 of the symptom-tracking applications ask at least one question about how the user is feeling, with the most common question being around feelings of anxiety. Given the potential effects that lockdown might be having on people's mental health, questions on this topic could be useful to mental health researchers to produce research insights that can be used to support and promote positive mental health and well being.

Outputs

Given the range of stakeholders involved in the symptom trackers examined, it is no surprise that there are a variety of outputs that are being produced. Outputs seem to fall into three rough categories:

Individual guidance

Once a user has filled out the questionnaire, they are presented with a set of recommendations. Many of the applications, such as the [CDC Symptom Checker](#), provide simple recommendations such as to see a health specialist if the user records the presence of Covid-19 symptoms.

Research insights

Many of the initiatives are producing research insights. Some of these insights are shared with the user, like [Your.MD's tracker](#), which gives a condition comparison after filling out the questionnaire, based on how you compare to other respondents. Other initiatives are sharing insights publicly, like the [COVID Symptom Study](#), which is producing visualisations such as how Covid-19 cases in the UK are changing day by day, or how they have changed specifically in the lockdown period. A number of the initiatives which are producing insights are also sharing data with health services, such as the NHS. [Project Oasis](#), for example, is a data intermediary which collects data from eight third-party applications and shares it with the NHS to support the response to Covid-19 in the UK.

Improvements to existing services

Some trackers, for example those that are part of wider health-related applications, are collecting Covid-19 data to improve their existing services, such as the support they can offer to individual users. These services aim to support users to improve their general health, rather than to specifically address Covid-19.

How does the data collected relate to the outputs?

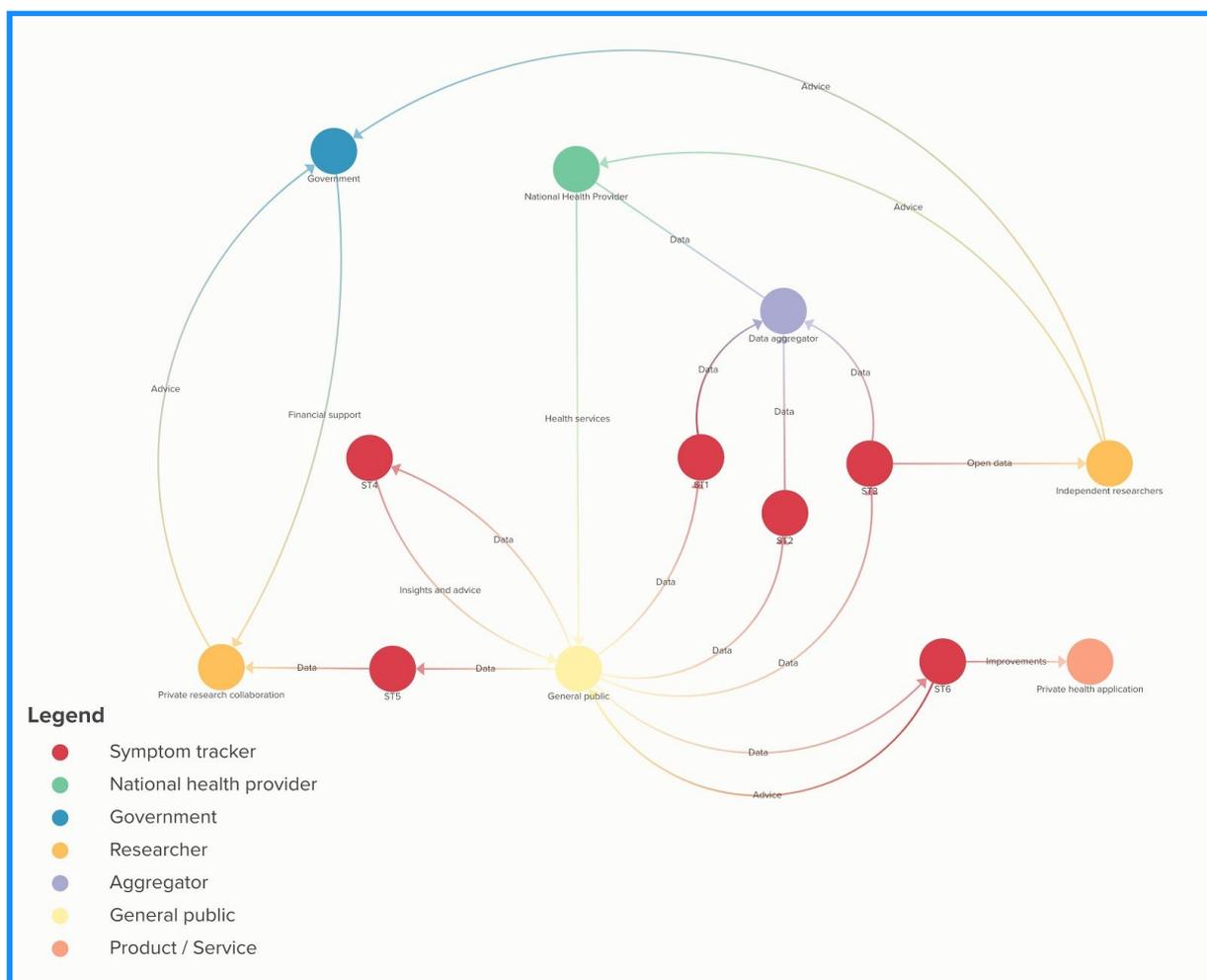
The level of detailed data required from users seems to differ greatly, depending on the outcome that the symptom trackers are trying to provide. For example, the [Ink C-19 app](#) requires only your location and an indication of whether you are well, have symptoms, have been diagnosed as sick, or have recovered from Covid-19. Once you have reported how you are feeling, a coloured circle representing your response appears on a map covering roughly a half-mile radius around your location, so other users of the application know that someone within that rough area has reported their status.

In comparison, the [Evergreen Life app](#) – which operates as a general health and wellbeing application, but has now also integrated a Covid-19 check into its functionality – asks for more detailed information, such as your current symptoms, whether you have been tested, whether you have been in contact with someone who has tested positive, and whether you would like the information you have submitted to be shared for NHS research purposes. This approach has yielded [outcomes](#) such as visualisations of how the number of people reporting Covid symptoms has changed over time, or heatmaps of how many people in a particular area are following existing government guidelines.

The symptom-tracking data ecosystem

Once we had completed the landscape review and data discovery exercises, we created two basic ecosystem maps to reflect what we had learned about the symptom-tracking data ecosystem.

The UK symptom-tracking data ecosystem



Example map of generic actors in the UK symptom tracking data ecosystem

In the UK alone, we identified 11 applications that could be considered symptom trackers. The majority of these initiatives were initially established as standalone collaborations between a few organisations. In May 2020, [jHub](#), the Innovation centre for Strategic Command, were tasked with identifying Covid-19 symptom trackers and helping them to transfer relevant, anonymous data to [NHSX](#). This project is called [Project Oasis](#).

Eight of the applications we identified have been verified as third-party suppliers by the NHS, and are sharing data as part of Project Oasis. These initiatives are:

1. [Agitate Ink C-19](#)
2. [Connected Cognition](#)
3. [Corona-Help UK](#)
4. [Evergreen Life](#)
5. [LetsBeatCovid-19](#)
6. [TrackTogether](#)
7. [Your.MD](#)
8. [C-19 COVID Symptom Study](#)

In addition, the NHS has its own online symptom checker which is available at [NHS 111 online](#).

Where is value being created in this example?

The eight third party suppliers that are part of the Project Oasis initiative are sharing data to jHub. jHub cleans and anonymises the data, then shares it with NHSX. This data is being used by NHSX to inform their COVID-19 response by helping them to understand the spread of the virus across the UK.

Each of these third party suppliers, and the other three initiatives – [CEBM COVID-19 Symptoms Tracker](#), [Patient Access](#) and [Studybugs](#) – are also using the data to create their own outputs, as described above.

Barriers to impact

One of the potential issues in this ecosystem is that while many of the third party applications are sharing data with NHSX via Project Oasis, and presumably with stakeholders involved in their own initiatives, there is very little transparency in what data has been collected, what is being shared between different stakeholders, and for what purpose.

The lack of visibility of what data is being collected could be stifling opportunities for collaborations between initiatives or with the research community more broadly. For example, researchers might have the skills or funding to try and address a particular research question, but might need access to data to help them do so. Without being able to see who is collecting useful data, researchers won't know which organisations to approach about collaboration.

The lack of transparency around how data is being used and shared also makes it very difficult for a user of the application to know what data about them is being used for and therefore make an informed decision about opting in or out of the service.

Opportunities

Symptom trackers in the UK ecosystem could provide opportunities for other researchers or innovators by being more open about the data that they are collecting. They could make descriptions of the data or metadata openly available on their websites, or through their applications, and by publishing open data where possible. This behaviour change could be driven by the funders of the initiatives, who might be able to spend money more efficiently by encouraging collaboration, or by the initiatives themselves, in order to solve any challenges they face by bringing in additional support. Initiatives like TrackTogether and Studybugs are currently leading the way in creating an open culture around Covid-19 symptom data and are an example for others to follow.

International variations

This section outlines what we have seen happening in other countries. Given the limited range of applications that were identified outside the UK, this is not a comprehensive review of the international approach to Covid-19 symptom tracking.

A single source approach

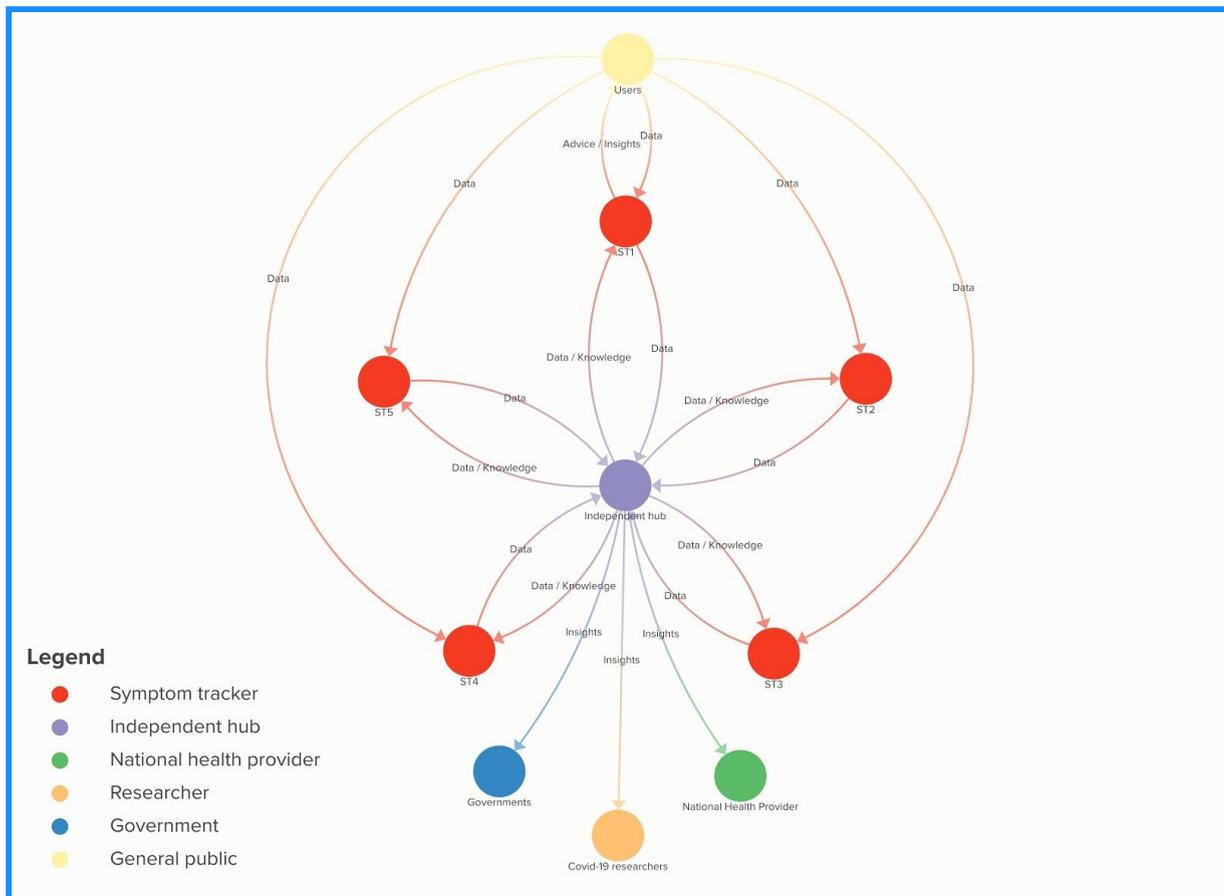
In some countries, citizens have been encouraged to use one particular application.

In Canada, the [Canada Covid-19 App](#) has been endorsed by Health Canada as the main application for Canadian citizens to use when checking symptoms or getting information about Covid-19. This allows the Canadian Government more control over how data is being collected about the pandemic, as well as an opportunity to disseminate information to citizens.

In Vietnam, a very similar approach has been taken with the [NCOVI app](#), the official application of the Ministry of Health and the Ministry of Information and Communications to help people assess their symptoms and make decisions about seeing a health professional. The application is also used as an official channel for state agencies to send recommendations to the public about national and local changes.

It is still likely that even if citizens are encouraged to use a particular application, that they could be using one or more alternative regional, national or international applications. This would likely be due to how outputs vary across applications. For example, parents of schoolchildren might find more value in an application that gives them tailored advice about their child and school, rather than one that recommends a national guidance.

A more collaborative approach



Example map of generic actors in a more collaborative symptom tracking data ecosystem

During this research, we were unable to find many good examples of where multiple symptom-tracking initiatives were working together to make data more accessible and produce insights about Covid-19.

One example that came up was the model adopted by the [Coronavirus Census Collective](#), an international consortium calling for academics, health organisations, and medical professionals around the world to join them. The Coronavirus Census Collective recognises that in this unique situation, a centralised platform could provide a place for different symptom-tracking initiatives to connect and learn from each other's approaches and findings.

Where is value being created in this example?

As well as enabling knowledge sharing, the platform provides a collection, aggregation and analytics function, which allows organisations to contribute to a growing data resource. The resource can be used by all members of the initiative. This approach enables people who need to access data about Covid-19 symptoms to find what they are looking for more easily.

A more collaborative approach can also help to address some of the inconsistencies around standards for data that might prevent data reusers from using symptom data to create new solutions.

Although the Coronavirus Census Collective data is being stored centrally, the outputs created by researchers and health professionals can still be tailored to their unique audiences.

Discussion

It has been good to see a large number of symptom-tracking initiatives emerging in response to the Covid-19 pandemic. With multiple organisations aiming to address this global challenge, we are seeing a variety of useful insights, from information about [local Covid-19 infection rates](#) to better understanding of [how the virus affects different people](#), being produced as a result of their efforts.

We believe that symptom trackers could unlock even more value from the data that they are collecting by collaborating with each other and sharing data. Better collaboration could provide data reusers with the opportunity to create new tools, guidance and approaches that researchers may not have the skills, expertise or resources to create themselves. However, our research has identified some major tensions that prevent collaboration and sharing of symptom-tracking data.

Data discovery

A key finding from this research is that the organisations behind symptom-tracking initiatives have not been very good at publishing information about the data that they are collecting. This limits the ability of data reusers to identify which organisations have useful datasets, and therefore who they should be approaching for collaboration.

Legal complexities

Negotiating legal terms for access to data can be complex – see the ODI’s [‘Data trusts in 2020’](#) blog post for more detail. Solving the complexities around negotiating access to data can incur a high cost and effort for both symptom trackers and data reusers who want to collaborate.

Technical barriers

An absence of compatible sharing mechanisms makes it difficult for symptom trackers to share data with reusers, without investing time and money into enablement. A lack of data standards around symptom-tracking data means that even if data is shared, significant effort might be required by reusers before the data can be used to best effect. Authorities have been stressing [the need for standardisation of Covid-19 data](#) and some data standards do already exist (for example, most of the data that has been shared has been through json and csv formats), but more effort is needed to create new standards or adopt existing ones, such as the [National Health Service SNOMED CT](#) standard for clinical vocabulary use in electronic records.

These barriers to sharing and collaboration could create unintended consequences for beneficiaries of symptom-tracking outputs. For example, unless they can access data from multiple symptom-tracking applications, data reusers can only draw data from a small portion of the target population. This could lead to outputs that don't reflect the diversity in our communities, and potentially have a negative effect on users – for instance generalising, and therefore misrepresenting, the risks that certain symptoms present to individuals.

Trust and trustworthiness

The pandemic has provided a number of health and wellbeing initiatives with a unique opportunity to collect useful data about people, but the lack of clarity about why some questions – such as personal changes during lockdown – are being asked is a matter of concern. It is clear from the research that symptom-tracking initiatives are collecting symptom data for multiple purposes, including providing personalised guidance and improving research insights about the virus. But they might also be using this as an opportunity to improve existing products and services, or to build new ones.

This in itself isn't necessarily a bad thing, but much of the data that is being collected is sensitive health data about people; this makes it important for symptom trackers to collect and use data in a trustworthy and transparent way, so that they don't unintentionally create harmful impacts for users. Our research has suggested that most symptom trackers could do a better job of communicating their purpose clearly and talking more openly about why they are collecting the information that they do.

Recommendations

We want public health officials, medical researchers and other data reusers to be able to get access to and reuse data collected by symptom-tracking applications, so that they can create new tools and make better decisions about the Covid-19 pandemic.

Activities involved in enabling better access and reuse of symptom-tracking data include:

- reducing the costs to application developers in providing symptom-tracking data
- reducing the costs to data reusers in accessing and using symptom-tracking data
- improving the understanding of application developers about the data needs of data reusers (for example which symptoms or other information needs to be collected to support operations or research, and the required frequency or regularity of this data collection)
- facilitating the discovery and negotiation of access to data between symptom-tracking applications and data reusers
- enabling data reusers to accurately assess the quality, utility and trustworthiness of data from symptom-tracking applications based on factors such as coverage, how users are asked for information, whether they have good ethical practices, and so on.

We believe that these outcomes can be facilitated through the creation of:

- **standards for sharing data** from symptom-tracking applications
- **an enabling hub to facilitate interactions** between symptom-tracking applications and data reusers, possibly including an audit and certification function.

We also believe that it is important for the symptom-tracking initiatives themselves to be open and transparent about what data they are collecting (at a high level) and why they are collecting, in order to build trust with the wider public.

Standards for data

Here we have summarised our thoughts on the importance of good standards for Covid-19 data. For a more detailed discussion, please read our blog post [‘Data and Covid-19: why standards matter’](#).

The ODI offers a free resource: the [‘Open standards for data guidebook’](#), which can help data stewards to understand the different types of standards, and provides guidance on when to create standards – and when not to.

Symptom trackers can create more interoperability without the need to start full-blown standardisation efforts. They can

- check whether someone else has already published similar data, and whether they can follow a similar approach
- if there is no precedent to follow, create clear and open documentation about the data they hold and how it is being structured, so others can follow the same approach.

This can provide a blueprint to enable new initiatives to follow in successful footsteps, or support existing initiatives to adopt a similar structure to the data that they have collected, making it more compatible with a successful approach. It will also help reusers of the data to understand what shape the data is in before they attempt to use it themselves.

Making adapters and other pipeline software open is another low effort, but useful way to increase interoperability in practice. For more information about the importance of standards and adapters for Covid-19 symptom-tracking data, please read our [‘Symptom tracking: standards and adapters’](#) guidance.

Governments should:

- In a time of crisis like this one, take an active role encouraging and helping increase interoperability by supporting collaboration on emerging standards
- When government-funded organisations are acting as aggregators, they should whenever possible make the adapters and software they are using as open as possible, so that others can benefit from data from different sources
- Outside of times of crisis, support standards infrastructure such as open standards organisations, so that standards can be available when needed and so that guidance can be available

Enabling hub to facilitate interactions

Whether you want to allow and nurture a rich decentralised ecosystem or champion a centralised approach, enabling the ecosystem of symptom trackers, data reusers and perhaps even beneficiaries to communicate and work together, is integral to making sure that all stakeholders are gaining collective value from data.

The development of a hub could help to make Covid-19 data more discoverable, reduce coordination costs associated with negotiating appropriate access to data between symptom-tracking applications and data reusers and help foster the exchange of knowledge.

Philanthropic funders, international public health organisations and government could/should play a stronger role in convening, funding these hubs, and maintaining knowledge there.

The ODI has developed the [Data Ecosystem Mapping tool](#), which can be used to understand who is involved in the ecosystem, identify where data is being collected and shared, where other exchanges of value are occurring, and to understand the impacts of these interactions. We recommend that health policy organisations use this tool to understand the current state of the Covid-19 symptom-tracking ecosystem and then use that understanding to consider what infrastructure needs to be built in order to enable the ecosystem to connect and collaborate more effectively.

Building trust with the public through ethical approaches

Symptom trackers need to make sure that as well as creating valuable outputs, they are protecting the rights and wellbeing of the people they are collecting data about, by being open and transparent about what data they are collecting and why, and using that data in an ethical way.

Data ethics relates to good practice around how data is collected, used and shared. It is especially relevant when data activities have the potential to affect people and society, directly or indirectly. The [Data Ethics Canvas](#) is a tool developed by us at the Open Data Institute (ODI) that can help those collecting, sharing and using data identify and manage ethical issues.

We recommend that symptom-tracking initiatives use the tool to:

- understand the data that they need to collect for their purpose
- explore the potential beneficial and harmful impacts of using this data
- plan their engagement with stakeholders, including identified beneficiaries
- integrate ethical practices into their project or initiative.

We have discussed the importance of identifying and managing ethical issues around data during the pandemic in [this blog post](#).

Appendices

Supporting documents

- [Research spreadsheet](#), which includes:
 - A longlist of symptom tracking initiatives
 - The shortlisting process for data discovery
 - A breakdown of the types of questions asked by organisation type