

## IDSEQ DATA PRIVACY NOTICE

Last Updated: October 10, 2018

IDseq is a cloud-based research tool used to make data-driven decisions about and to support global public health and the study of infectious diseases and pathogens. To do this, IDseq processes genetic data in order to identify pathogens contained within.

The Chan Zuckerberg Biohub Inc. (“CZ Biohub,” “we,” “us,” or “our”) provides IDseq (“Services” or IDseq). This Data Privacy Notice describes the types of information we collect from website visitors and registered users (“Users”), and how we use, share, and protect that information IDseq.

Here’s how IDseq works: Users upload Research Data. This data contains human and non-human genetic sequences, as well as information about those sequences, such as the date the sample was collected and the species it was collected from. For example, a researcher might upload genetic information from mosquitoes, which are often a source of infectious disease or from humans, who can be infected by such diseases. IDseq then processes this Research Data in order to identify pathogens found within the genetic sequence (e.g., the mosquito may be carrying the chikungunya virus). IDseq also collects information about visitors and Users in order to offer the Service. Other than basic information required to create an account (e.g. email address, name), the User determines what information they want to upload onto IDseq.

### **I. Research Data**

Research Data is data that Users upload to IDseq. As explained below, Research Data consists of genetic sequence information (human or non-human) and non-personally identifying metadata about those genetic sequences (such as time and location of sample collection).

As described in our <Terms of Use>, Users are required to obtain and maintain all necessary consents, permissions, and authorizations required by applicable laws needed prior to uploading, sharing, and exporting Research Data with the Services.

### **What Is Research Data?**

Research Data includes Sample Data, Sample Metadata, and Report Data.

- ***Sample Data.*** Sample Data is full genetic sequence data uploaded by Users. Genetic sequence data is machine-readable DNA code data containing the genetic sequence of the host from which the sample was taken. The host could be a human or non-human (e.g., mosquito).

- **Sample Metadata.** Sample Metadata includes information related to the Sample Data, such as the host type (e.g., human or mosquito), upload date, location, tissue type, and free-text research notes entered by Users. This data should not include personally-identifying information regarding the individual to whom the Sample Data relates.
- **Report Data.** Report Data is information IDseq produces from the Sample Data and Metadata. To produce Report Data, IDseq uses data cleaning (e.g., removing duplicate nucleotides) and analysis (e.g. matching nucleotide sequences in the Sample Data to known pathogen sequences) – a process commonly known as a Data Pipeline. Report Data may include, for example, data about the pathogen sequences identified in the Sample Data and the frequency of such identification (“Pathogen Data”) or raw numeric counts of non-personally identifying protein expression profiles that were found in the Sample Data (“Gene Counts”). Once Sample Data has been put through the Data Pipeline, the Report Data that is produced no longer includes full genetic sequence data, is not personal information, and does not, on its own, permit direct association with any specific individual.

### **How We Use Research Data**

Research Data is used for the following purposes:

- To create Report Data, including new reports for Users when we update our Data Pipeline.
- To improve the way IDseq creates Report Data, including improving our Data Pipeline.

We never sell Research Data.

### **How We Share Research Data**

Sample Data and Sample Metadata are shared with the Users that uploaded the data, as well as other IDseq users they expressly share that information with. Users can share the Research Data they upload by creating Projects (groups of Samples and Reports) and sharing those Projects with other IDseq users, who will then be able to see and download the Research Data that was shared with them. Research Data is also shared as described in the section on Partners, Vendors, and Other Third Parties below.

Report Data is visible to all IDseq Users, and they may share it with others. We hope that this sharing of pathogen data will help to create a global dashboard that helps give researchers, for example, early warning into the next outbreak or virus.

### **What’s our legal basis to use and share Research Data?**

To the extent that the European Union’s General Data Protection Regulation (“**GDPR**”) applies, we rely on the following legal bases to use and share Research Data:

- The explicit consent of the individual whose data is being analyzed, where such consent has been obtained by the User in accordance with the GDPR; and
- The public interest and our and our Users’ legitimate interest in investigating and stopping the spread of infectious diseases and promoting global health. The use and sharing of Research Data furthers the public interest in the area of public health, particularly by helping to protect against serious cross-border threats to health. The processing of Research Data is also necessary for scientific research purposes.

## **II. Visitor and User Data**

Visitor and User Data is the information we collect from your use of IDseq.

### **What We Collect**

**Visitor Data.** This is information collected from visitors to our website, whether or not they are Users (“Visitor Data”). When you visit our website, we may automatically collect Device Data or Analytics Information. Device Data may include information about your browser type and operating system, IP address and/or device ID, including basic analytics from your device or browser. Analytics Information relates to any of your requests, queries, or use of the Services, such as the amount of time spent viewing particular web pages. We use Google Analytics for this service. Visitor Data may also include information we collect about you when you visit our website or use the Services, including through the use of cookies, web beacons, and other technologies.

- **Cookies, Web Beacons, and other Technologies.** A cookie is a small file that may be stored on your computer or other device. Web beacons and similar technologies are small bits of code are embedded in web pages, ads, and email that communicate with third parties. We use these technologies to recognize your device and browser and do things such as keep you logged in or to understand usage patterns.

**User Data.** User Data is any information we collect from an IDseq User about that User, including Visitor Data. It may include information necessary to create or access your account such as your name, username, email address, and login credentials.

### **How We Use That Data**

Visitor and User Data will be used for the following purposes:

- To identify you, create a profile for you, and verify your identity so you can login to and use IDseq.
- To provide you with notices about your account and updates about the IDSeq services.
- To respond to your inquiries and requests.
- To analyze how Users are using IDseq so we can improve IDseq.
- To protect the security and integrity of the Services.

### **What's our legal basis for using Visitor and User Data?**

To the extent that the European Union's General Data Protection Regulation ("GDPR") applies, we have a legitimate interest in using Visitor and User Data in the routine ways described above to provide, protect and improve the Services. This allows us to improve the Service we provide to users which, in turn, supports research regarding global public health and the study of infectious diseases.

### **III. Partners, Vendors, and Other Third Parties**

The Chan Zuckerberg Initiative, LLC ("CZI") is our primary partner in building IDseq. CZI helps design, build, and operate the Service. As a result, all data in IDseq is shared with CZI and it is a joint data controller (under EU privacy law) of all personal data in the IDseq Service. CZI processes data in accordance with this Privacy Notice and access to any data within IDseq by CZI is limited to those staff who need access to operate the service.

We also share Research, Visitor, and User Data with third parties who provide services to us or CZI, such as database providers like Amazon Web Services. We may also share this data with analytics vendors that assist us to improve and optimize our Services.

We may share Research, Visitor, and User Data with any third party that purchases, or to which we transfer, all or a portion of our assets and organization. Should such a sale or transfer occur, we will use reasonable efforts to try to ensure that the entity to which we transfer your personal information uses it in a manner that is consistent with this Data Privacy Notice.

We may disclose Research, Visitor, or User Data if we believe in good faith that such disclosure is necessary (a) in connection with any legal investigation; (b) to comply with relevant laws or to respond to subpoenas or warrants served on us; (c) to protect or defend our rights or property or those of users of our Services; and/or (d) to investigate or assist in preventing any violation or potential violation of the law, this Privacy Policy, or our <Terms of Use>.

#### **IV. How We Protect the Information**

IDseq uses industry standard security measures to ensure the confidentiality, integrity and availability of data uploaded into the platform. This includes best practices like encrypting connections to IDseq using TLS, hosting IDseq on leading cloud providers with robust physical security, and ensuring that access to any data within IDseq by CZI and Biohub staff is limited to those staff who need access to operate the service.

Security takes ongoing work and we will continue to improve our security measures as IDSeq develops. Please notify us immediately at [security@idseq.net](mailto:security@idseq.net) if you suspect your account has been compromised or are aware of any other security issues relating to IDseq.

IDseq is not designed for or directed toward children under the age of thirteen.

#### **V. How Long We Retain Data and Data Deletion**

Research Data is retained until Users delete it from IDseq. Users may submit deletion requests by emailing [privacy@idseq.net](mailto:privacy@idseq.net) and we will delete the requested Sample Data and Sample Metadata within 60 days. Please note that improvements to our Data Pipeline from processing the Research Data cannot be removed.

User Data is retained until Users delete their IDseq account. Users may submit account deletion requests by emailing [privacy@idseq.net](mailto:privacy@idseq.net).

#### **VI. Choices About Your Data**

If you are a User, you have the following choices:

- Users are able to request the deletion of User Data that constitutes your personal information or Research Data that you uploaded to IDseq.
- Users are able to access and download their Report Data within IDseq.
- Users may also object to the processing of Research Data in certain circumstances by emailing [privacy@idseq.net](mailto:privacy@idseq.net). In such cases, we will stop processing that data unless we have compelling legitimate grounds to continue processing or where it is needed for legal reasons.

If a User has uploaded Research Data containing data about you, please see below:

- We require Users who upload Research Data to ensure they have all necessary consents, permissions, and authorizations to do so. We are not in a position to confirm which data uploaded by Users may relate to other individuals. As a result, currently IDseq is able to receive access, restriction, rectification, objection or deletion requests only from Users.

- If you believe your information has been uploaded to IDseq, you should contact the researcher or User that uploaded this information to (i) request access to the information, (ii) object to the processing of the information or (iii) seek deletion, restriction or rectification of the information. Similarly, where someone has previously provided consent to a researcher or User, that person has the right to withdraw that consent and should contact the User to make such withdrawal or otherwise register an objection to the use of their data.

Please contact us at [privacy@idseq.net](mailto:privacy@idseq.net) if you would like to exercise the privacy choices discussed above or have any questions. If your data is subject to the GDPR and you wish to raise a concern about our use of your information (and without prejudice to any other rights you may have), you have the right to do so with your local supervisory authority.

## **VII. Data Transfers**

IDseq is a global service. By using IDseq, Users authorize us to transfer and store the uploaded information outside your home country, including in the United States, for the purposes described in this Privacy Policy.

If you want to use the Services, you need to agree to our **<Terms of Use>**, which set out the contract between IDseq and our users. As we operate in countries worldwide (including in the United States) and use technical infrastructure in the United States to deliver the Services to you, in accordance with the contract between us, we need to transfer your personal information to the United States and to other jurisdictions as necessary to provide the Services. Please note that the privacy protections and the rights of authorities to access your information in these countries may not be the same as in your home country.

## **VIII. How to Contact Us**

If you have any questions, comments, or concerns with this Data Privacy Notice, you may contact our DPO by email at [privacy@idseq.net](mailto:privacy@idseq.net) or by physical mail at:

Chan Zuckerberg Biohub 499 Illinois Street, Fourth Floor San Francisco, CA 94158	Chan Zuckerberg Initiative, LLC P.O. Box 8040 Redwood City, CA 94063
--	--

## **IX. Changes to This Data Privacy Notice**

This Data Privacy Notice was last updated on October 10, 2018. We may update this Data Privacy Notice from time to time and will provide you appropriate notice of any updates.