

## Interruption to GISAID-based SARS-CoV-2 sequence analyses – GISAID Fact Check

Trevor Bedford, Richard Neher and the Nextstrain team

2025-11-06

This document represents an addendum to the Nextstrain blog post:

[nextstrain.org/blog/2025-11-06-gisaid-based-ncov-analyses](https://nextstrain.org/blog/2025-11-06-gisaid-based-ncov-analyses)

Below is our fact check outreach to GISAID on Mon Nov 3 in plain text, *verbatim responses from GISAID on Wed Nov 5 in italic text* and **our clarifications on Thur Nov 6 in bold text**.

Subject: Fact check & on-the-record comment re: forthcoming Nextstrain post

Dear GISAID Secretariat,

We feel we need to inform our users as to the situation with SARS-CoV-2 analysis updates. We plan to publish a short status update to [nextstrain.org/blog](https://nextstrain.org/blog) about GISAID-dependent SARS-CoV-2 analyses on Wednesday Nov 5. For accuracy and fairness, please confirm/correct the brief facts below and share an on-the-record statement (150–250 words).

Facts we plan to reference

(1) On Oct 1, 2025, GISAID informed us that updates to the Nextstrain-specific flat-file feed would be discontinued effective immediately.

*GISAID: It is important that you inform your users that GISAID only provides access to natural persons, not groups, institutions or entities. GISAID's agreement, and the terms of use of the GISAID platform, were with you (Trevor Bedford), not a group, institution or entity. We thus believe the "us" is not accurate.*

**Nextstrain: The original blog post has been revised to clarify that the flat file was provided to Trevor Bedford the individual and not to Nextstrain the entity.**

*GISAID: In addition, this statement is incomplete. We request that you provide the specific context regarding the discontinuance that was explained to you: your 'nextregion' package is not being used in a manner and a frequency that justifies the effort required by GISAID to package and prepare the information. In particular, the specific reason provided to you is that over a seven-month period merely 79 users made a total of 88 downloads of your package.*

**Nextstrain: We've updated the original blog post to include this specific detail about stated rationale being downloads of the 'nextregion' package.**

(2) GISAID proposed that we provide Augur parameters, GISAID would run Augur, and return subsampled Auspice JSONs for our trees.

*GISAID: Again, this statement is incomplete and misleading. GISAID stated, and continues to state, that it is happy to continue to support your work to keep up-to-date your global subsampled phylogenetic tree of the 'Genomic epidemiology of novel coronavirus (nCoV)'. In addition, your statement should note that this was the only tree GISAID had agreed to support in January 2020 (see your email update 26-Jan-2020). Again, as the terms of use and agreements of GISAID run to individuals, you should not imply that GISAID has entered into any sort of agreement with a any group, institution or entity group, institution or an entity (legal or otherwise). These requirements (agreements with individuals, not institutions) were requested by the individual submitters when GISAID was first created.*

**Nextstrain: We find this characterization odd as correspondence from GISAID Secretariat states: "Once you provide GISAID with all necessary criteria for set selection, subsampling and target size for each tree. GISAID technical staff can gladly look into the modalities of enabling the 56 trees you are making reference to."**

(3) GISAID stated that packaging/transferring large flat files is not sustainable.

*GISAID: Again, this statement is incomplete. As noted above, the reason for GISAID's discontinuance was clearly explained to you. GISAID regularly evaluates efficient usage of resources, and indicated to you that because the usage of your 'nextregion' package was so low (i.e., during the past seven months (Mar-Sep 2025) only 79 users have made a total of 88 downloads), GISAID could no longer support the work required for such low usage.*

**Nextstrain: We believe the focus on 'nextregion' downloads from within GISAID's EpiCoV interface to be an intentional red herring and obfuscation. The primary public health utility of Nextstrain analyses is a public dashboard for keeping tabs on continuing SARS-CoV-2 evolution. This dashboard at [nextstrain.org/ncov/gisaid](https://nextstrain.org/ncov/gisaid) has seen >36,000 unique visitors with >71,000 pageviews in the seven months from March to September 2025. Making subsampled sequences and metadata available through 'nextregion' was a simple byproduct of not being allowed to reshare data directly.**

**Additionally, the timeline for the flat-file data feed is as follows. On Jan 19, 2020, we began updating [nextstrain.org/ncov](https://nextstrain.org/ncov) from manually downloaded sequences from GISAID. On Jan 30, 2020, GISAID provisioned the flat file for updating [nextstrain.org/ncov](https://nextstrain.org/ncov). Starting May 13, 2020, we provisioned files for GISAID to rehost, which they subsequently made available under the heading of 'nextregion'. We've included screenshots from Jan 30 and May 13 emails as evidence of this timeline.**



**Sebastian Maurer-Stroh**

Privileged access for automatic updates

To: Trevor Bedford, Cc: Peter Bogner

January 30, 2020 at 2:45 AM



[Details](#)

Dear Trevor,

the technical team has managed to set up a way for you to access full-dumps (meta + sequence) at:  
[https://platform.gisaid.org/epi3/3p/nextflu/\[REDACTED\].json.bz2](https://platform.gisaid.org/epi3/3p/nextflu/[REDACTED].json.bz2)

It is updated with the live-data every 15 minutes.

You can access it as user "tbedford" with [REDACTED] credentials.

It is a simple linewise-JSON-File. Attached a Document describing the field names.

Needless to say that this is a great sign of trust and privileged treatment to foster your partnership with GISAID and I hope you appreciate the magnitude of this gesture towards you.

Best regards,  
Sebastian



**Trevor Bedford**

Re: ncov.json CANNOT be accessible as it is now

To: Sebastian Maurer-Stroh, Cc: Gunter Bach, Peter Bogner

May 13, 2020 at 11:30 AM

[Details](#)

Hi Gunter,

Very sorry about that. I just added the bucket "nextstrain-data" to the "GISAID" user credentials. You should now be able to programmatically access files in the "nextstrain-data" bucket with your existing credentials. Can you check if this works and let me know if it does not?

- Trevor

On May 13, 2020, at 9:01 AM, Sebastian Maurer-Stroh <[sebastianms@bii.a-star.edu.sg](mailto:sebastianms@bii.a-star.edu.sg)> wrote:

Hi again,

Gunter does not have the credentials for this S3-Repo.  
He only has the credentials for "[s3://nextstrain-ncov-private/](#)"  
Can you please provide to him directly (will go to sleep at some point)?

Cheers,  
Sebastian

Dr. Sebastian Maurer-Stroh :: Deputy Executive Director (Research) :: Bioinformatics Institute (BII) :: Agency for Science, Technology and Research (A\*STAR)  
30 Biopolis Street #07-01 Matrix Singapore 138671 :: DID (65)64788377 :: Fax (65) 64789047 :: <http://www.a-star.edu.sg>

On Wed, May 13, 2020 at 11:08 PM Trevor Bedford <[trevor@bedford.io](mailto:trevor@bedford.io)> wrote:

Hi Sebastian,

These files are available on S3 as:

- [s3://nextstrain-data/ncov\\_global.json](#)
- [s3://nextstrain-data/ncov\\_africa.json](#)
- [s3://nextstrain-data/ncov\\_asia.json](#)
- [s3://nextstrain-data/ncov\\_europe.json](#)
- [s3://nextstrain-data/ncov\\_oceania.json](#)
- [s3://nextstrain-data/ncov\\_north-america.json](#)
- [s3://nextstrain-data/ncov\\_south-america.json](#)

(4) Since October, we have requested parity access to the same non-Nextstrain-specific flat files provided to other external platforms; if legitimately not possible, we're open to the GISAID-run option wherein JSONs are shared through Nextstrain Groups interface.

*GISAID: It is important to remember that user "access" to data on the GISAID data sharing platform and having the privilege of receiving a custom data provision to enable a dashboard with data shared via GISAID, are two entirely different things. What you appear to be referring to are data provisions such as the one that had been provided to you to generate your 'nextregion' package. Therefore, the concept of parity that anyone is entitled to a data provision is unrealistic. Data provisions from GISAID, including those that enable dashboards, are currently under review. This is of particular importance as GISAID is aware, and addressing, the fact that there are several individuals that have misused a GISAID's data provision and their agreements. Among them are individuals that received GISAID support to operate their own "dashboard" to display aggregate data, but have abused the GISAID data sharing process in violation of the rights inherent to the owners of that data. See <https://gisaid.org/resources/statements-clarifications/data-related-complaints/>.*

**Nextstrain: We would still welcome an explanation as to why it is not possible to give us access to the same flat file that is being provisioned for operating external websites. We believe this decision is actively harming global genomic surveillance without sufficient cause.**

(5) In 2021, we implemented requested changes (retain hCoV-19/ prefix, updates to tooltips, no metadata downloads).

*GISAID: This statement is again incomplete, and neglects to state that the reason those changes were implemented were because of violations of the terms of use you agreed to when you signed up to use GISAID (e.g., by operating the now defunct influenza database (who.nextflu.org) that provided data from GISAID EpiFlu). Your release of metadata from GISAID EpiCoV on GitHub was also a violation of GISAID's terms of use, as well as enabling the download of FASTA sequence data from your global subsampled nCov tree, which ultimately lead to your call with Drs. Kieny, Fouchier and Maurer-Stroh on 15-June-2021. During this call, only compliance improvements to your global subsampled nCov tree that were immediately implementable were discussed.*

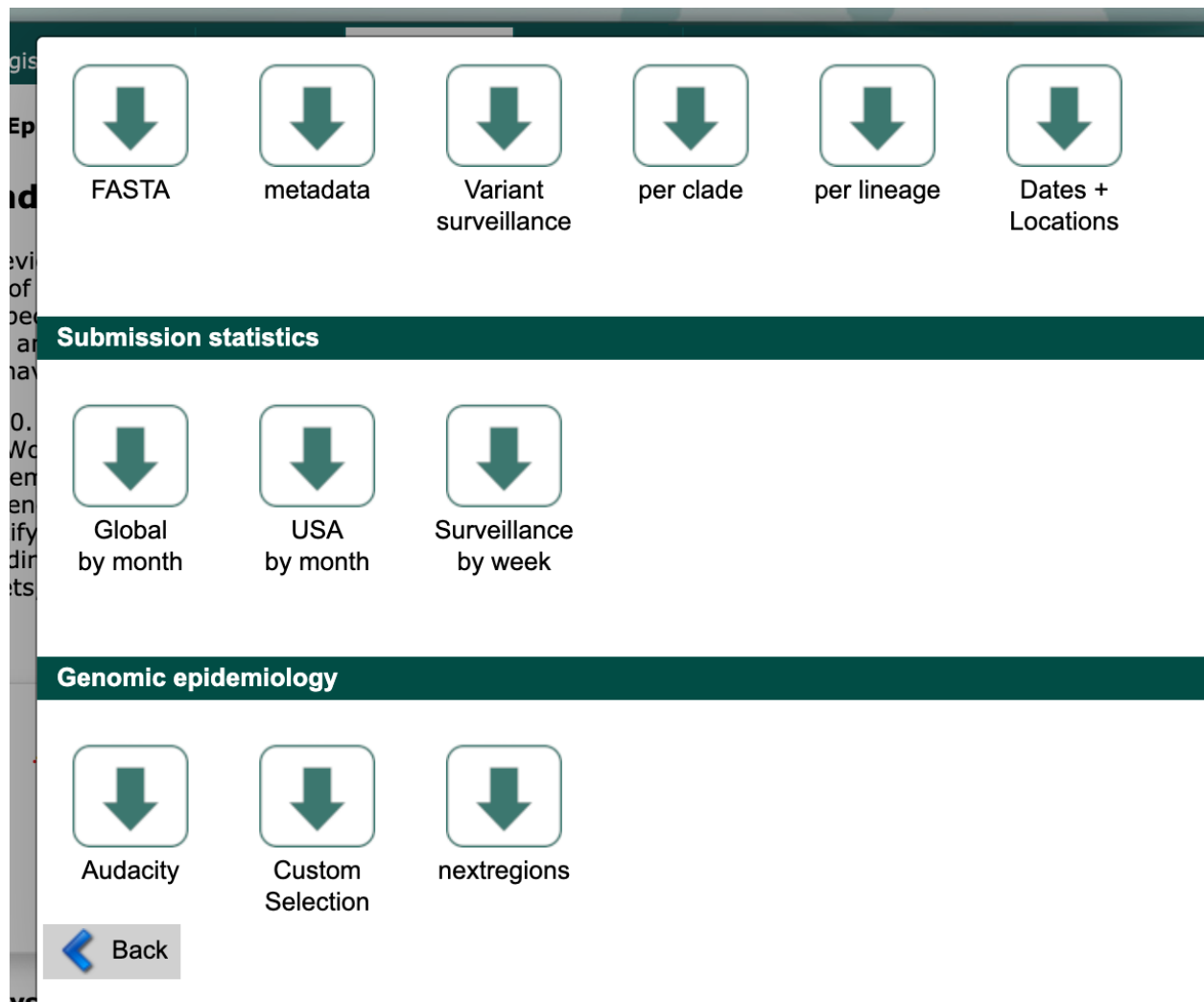
**Nextstrain: We believe much of this is a misrepresentation. who.nextflu.org was a private site viewable only by members of the WHO GISRS network and did not use sequences from GISAID in any different of a fashion from the public nextflu.org and nextstrain.org/flu analyses. We've never actually been told about GISAID's specific allegations regarding [who.nextflu.org](https://who.nextflu.org).**

**In March 2020, we had provisioned metadata including geographic location and sample collection date publicly to GitHub. This was due to confusion on the part of Trevor Bedford as to what data from GISAID was permissible to share. This stemmed from many papers including metadata in supplementary tables ([Data S1 in Forster et al](#)) or directly in**

phylogenetic figures ([Fig 1 of Su et al.](#)) It's also common practice to redact sequences but not metadata in effecting reproducible research for GISAID-based analyses ([BEAST XML](#) from [Worobey et al.](#)). I (Trevor) believed that similar conventions surrounded websites and as that surround papers. When informed that GISAID viewed sharing metadata as a violation of Terms of Use, we immediately complied and removed public metadata.

*GISAID: It is also important to tell your audience that you never consulted with, nor sought permission from GISAID, to make use of the data package provided to enable your nextregion package for other trees, leave alone for webservices. GISAID asked you for an explanation of how you used GISAID data as early as 2022, but you have yet to provide any response.*

**Nextstrain:** We believe this is also a misrepresentation. In the May 13 email above it's clear that from the beginning 'nextregion' was based on multiple trees. You can see this further in the screenshot of EpiCoV interface below that shows "Region-specific Auspice source files" for Africa, Asia, etc... under "nextregions".



## Region-specific Auspice source files



Global



Africa



Asia



Europe



North  
America



South  
America



Oceania

6. GISAID has asked other groups (not Nextstrain) for a ~\$100k annual fee for continued data access. We'd like to rebut online [rumors like this one](#).

*GISAID: You are aware that the statement in the X post is not true because GISAID never requested any money from Nextstrain. Thus, attempting to rebut "rumors" does nothing but further re-publish those same false statements. Moreover, the X post has been taken down once we informed the individual who made the post that the statements, allegations, and innuendos in the post were not correct.*

*To be clear, it is utterly false for anyone to suggest that GISAID has asked for a \$100k annual fee for continued data access. GISAID has never demanded or otherwise sought any money from CovSpectrum for updated data, let alone \$100,000. Similarly, GISAID has never demanded or otherwise sought any money from Outbreak for updated data, let alone \$100,000. And, as noted above, GISAID has also never demanded or otherwise sought any money from you for updated data, let alone \$100,000.*

*We thus ask that you not repeat or republish these incorrect and false "rumors," as GISAID has now informed you that they are not true.*

**Nextstrain: We've updated the original blog post to strike reference to annual fees requested to Nextstrain or others.**

**Optional clarification (brief answer welcome):**

A) Is a standard programmatic flat-file feed available to external analysis sites, and what is the request process?

*No response to this optional query.*

**Nextstrain: We would still welcome public engagement from GISAID on their decision making regarding external analysis sites and a clear request process for individuals who wish to build and maintain such sites. This is especially relevant with ending updates to Nextstrain, CoV-Spectrum and outbreak.info.**

Please send factual corrections and your statement by 9am PT Wednesday morning. We'll incorporate factual corrections and include your statement verbatim (edited only for length/format if over 250 words). If received later, we'll append it with a timestamp.

Sincerely,  
- Trevor