This is a supportive document for the <u>SimEx - Curation Team Feedback spreadsheet</u>. Some issues may be best described / documented using screenshots. If you would like to provide feedback in screenshot form, please use this shared document. Remember to write "SCREENSHOT" in the appropriate cell in the spreadsheet and our team will know to review this log.

Curator Log	Screenshot	Comment
<b>EXAMPLE</b> A.B.C. 2024-03-10	GENERAL ⑦ Case status	I identified a case status from a source not captured in our data dictionary. The case was described as "Probable." [link to source] I communicated this issue to the curation team in the Slack #ingestion channel and we collectively decided to log the case as "suspected". Do we have a defined process for how this works during a real outbreak?
J.S. 2024-03-14		some of the event dates that I didn't fill in were automatically filled in - bulk add button did not work Issue was identified upon review of submission and brought up in the slack #ingestion Noted that there are two spots for date when it is indicated that the case has traveled. One says travel history entry (which is vague) and the other says travel history start both seem a bit vague

	TRAVEL HISTORY ⑦		
	Travel history entry		
	Travel history start		
	Europe		
	Travel history country		
S.K. 2024-03-14	Case 1		Travel history was entered for case 1 and case 2 but this data does not appear in the download.
	Has travel history Travel history entry Has travel start Last known location Last known country	Y Nigeria	

	Id         travelHistory.travel.location.country         travelHist         travelHist	
S.K. 2024-03-14	_id       dat       events.dateConfirmation       e         1       2022-05-07T00:00:00.000Z       2       2022-05-18T00:00:00.000Z         3       2024-03-14T00:00:00.000Z       4	9 cases were added in one entry for the UK cases reported from the US new source (cases 3-11). events.dateConfirmation appears to be populated with the entry date, but only for the first case in the batch (case 3) (see screenshot).
SK 2024-03-14	id       events.realcocation.administrativeAreaLevel1       location.administrativeAreaLevel2       location.administrativeAreaLevel3       location.country         1       United Kingdom       United Kingdom       United States of Area         LOCATION       Geo resolution       Point       Docation       Docation         Admin1       England       Admin2       Greater London       Admin3       London         Location       London       Greater London       England       Admin3       London         Location       London       Greater London       England, United Kingdom       Lotation       Location         Location       Condon       Greater London       England, United Kingdom       Location       London, Greater London, England, United Kingdom         Latitude       51.5073       Congitude       -0.1276       Comment       Guy's and St Thomas' NHS Foundation Trust, London	Case 1 - admin levels for location do not appear in the downloaded data.
J.A.S. 2024-03-14	Image 1 Case 2 EDT VERIFY O CURATORS Created by Verified by Not verified Image 2	Image 1: No name or email is present to say who created the case and who verified it added to Github

	EVENTS ⑦ Entry date 2024/03/13	Image 2: Concerned that dates such as confirmed case date, onset of symptoms, and consultation date are being auto-populated if the curator does not enter this info in. added to Github	
	Reported date 2022/05/17 Confirmed case date 2024/03/14 Method of confirmation		
	Onset of symptoms date 2024/03/14		
	First clinical consultation date		
	Image 3: Cases by Country	Q Search X	Image 3: Pivot table totals account for omitted cases as well (see UK) added to
	United Kingdom 1 8 United States of America 1 Grand Total 2 8	10 1 11	Github
J.A.S 2024-03-18	Image 1		Image 1: column names are long (i.e., events.outcome) and confusing. For users who are downloading our data we may want simple names such as "Outcome" added to Github

	AD AE AF AG AH AI AJ
	events.iso events.out events.rea location.a location.a location.a locatio
A.H. 2024-03-14	Image 1: Curator's comment AH Edited Date first consultation; removed auto- populate throughout Image 2: versponse should NOT have a body Image 3: versponse should NOT have a body Image 4: versponse failed with status code 500 Sort by
H )24-03-19	Image 1:

	LOCATION 🕜			Admin 3 and Lat Long data points that were also incorrect. (Image 1) However, after I committed the changes and went back to check the updated case data,
	Location to geocode		•	
	Country * ♥ United States ▼	Admin 1	•	the location still read as 'Miami, Florida, United States.' (Image 2) There is no way to clear this data.
	Admin 2 Broward	Admin 3	•	
	Latitude	Longitude		
	Location Comment Image 2:			
	LOCATION Geo resolution Country Admin1 Admin2 Admin3 Location Latitude Longitude Comment	Point United States of Ar Florida Broward Miami, Florida, Unit undefined undefined		

AH 2024-03-19	CURATORS Created by Verified by	Data does not populate for who created and verified a case. added to Github
JAS 2024-03-25	Full Data Download         - Full dataset is just named "full" this is not a very descriptive title         Image 1: Too much source information            (effective source)	<ul> <li>Image 1: all the source information is spread across the sheet and not in one congruent location. In addition, all the "case reference" information is too much and not of interest to users downloading it. We really only need the primary source and additional sources referenced. added to Github</li> <li>Image 2: too much information (i.e., the time thing) is downloaded with the dates. One of the confirmation dates is attached to a suspected case (2024-03-14)! No date was entered for this date in the portal, it just autofilled in! We don't want that. added to Github</li> <li>Image 3: including latitude and longitude in the dataset feels like overkill. Most users probably care about country level information only. Additional location information that is helpful is state, county, province, and city if it's available.</li> <li>Image 4: Should this say the name of the</li> </ul>

	AL	AM	AN	pathogen we are tracking? In this case would
	location.geometry.la	location.geometry.longitud	location.geoResolutic	it say Mpox? added to Github
	51.507322	-0.127647	Point	
	42.32047006	-71.96032576	Point	Image 5: Users do not need our revision
	54.23793336	-2.36966957	Point	metadata - they don't need to know the exact
	54.23793336	-2.36966957	Point	date/time we made changes or edits, how
	54.23793336	-2.36966957	Point	many times we made revisions, etc. added to
	54.23793336	-2.36966957	Point	Github
	54 23793336	-2 36966957	Point	
	54 23793336       -2 36966957 Point         Image 4: Pathogens column data is blank         AQ       pathogens         pathogens		<ul> <li>I did not capture an image of these because I ran out of time but wanted to document these notes: <ol> <li>Too much travel history information is downloaded. It's similar to location information: we don't need to share the exact geolocation of the travel site(s)</li> <li>Not a fan of how the data download includes EVERY data element that might be included for a case, even if none of the cases have any of that info (i.e., we have no information for vaccines in this dataset yet all those blank columns downloaded and it makes the sheet look messy). For reference my downloaded data set has columns A-CH. This is not the same case for filtered data, only select data elements are downloaded.</li> </ol> </li> </ul>	
JAS 03/28/2024	Image 1			Image 1: Date last modified column is empty, added to Github



\*Create additional rows, as needed, to add multiple screenshots with associated comments.