

Public health taxonomy for social listening on mpox conversations

For use in infodemic monitoring and insights generation (updated on 20 September 2024)

Background

The infodemic has been defined as 'an overabundance of information, accurate or not, in the digital and physical space, accompanying an acute health event such as an outbreak or epidemic'¹. The first steps in managing an infodemic are social listening and integrated analysis. Integrated analysis is best done with a diverse set of data sources across social, program, behavioral, media, and epidemiological domains. Infodemic monitoring and intelligence gathering consider as many relevant data sources as possible, recognizing they all individually have limitations and biases, but when analyzed in an integrated manner, can yield insights that data sources analyzed in isolation would not provide.

A public health taxonomy is a system of classification or structure that allows an analyst to align data with a search strategy to better understand domains related to a public health topic of interest². Taxonomies are generally used to help analyze data that make up an infodemic insights report, which is used to inform public health action to better address people's questions, concerns, information voids, narratives and circulating mis- and disinformation, all of which make up the information environment and can affect people's risk perception, awareness and behavior on a health topic. A taxonomy can help organize and map the information on, perception of, discussion and behaviour around a public health topic that are most concentrated, where conversations seem to be missing or silent, and where new topics or concerns may emerge. Taxonomies have been developed and implemented by the World Health Organization (WHO) for respiratory pathogens, COVID-19, and for mpox. Infodemic monitoring and insights generation encompasses systematic **analysis of different data sources**, generation of **intelligence**, turning the intelligence into infodemic **insights** and applying structured judgment and risk matrix approach to generate **recommendations** for action.

Objective

Millions of daily conversations by members of various communities can provide a better understanding of how the multi-country outbreak of mpox outbreak is perceived, understood and discussed. The objective of the taxonomy is to support the monitoring of conversations related to mpox within thematic categories relevant to public health response. The subtopics are to capture breadth of conversations and help identify structure and changes in narratives. The taxonomy can be applied to both online and offline social listening and can facilitate structured integrated analysis of the data to produce infodemic insights and recommendations.

¹ Briand S, Hess S, Nguyen T, Purnat TD. Infodemic Management in the Twenty-First Century. In: Purnat TD, Nguyen T, Briand S, editors. Managing Infodemics in the 21st Century : Addressing New Public Health Challenges in the Information Ecosystem. Cham: Springer International Publishing; 2023. p. 1-16.

²Purnat TD, Nguyen T, Ishizumi A, Yau B, White B, Cecchini S, Samuel R, Hess S, Bezbaruah S, Briand S. 2022. Delivering actionable infodemic insights and recommendations for the COVID-19 pandemic response. Weekly Epidemiological Record. 2022 Jul 8; 97 (27), 313 - 324. <https://apps.who.int/iris/handle/10665/359144>

Notes to analyst:

When analysing data using a taxonomy, it can be useful to consider using a range of approaches to understanding and describing data.

- Compare the volume and themes of overall net conversations with mpox conversations.
- Compare the velocity of conversations within the mpox categories over time.
- Identify mpox-specific themes of questions, worries, concerns, information voids, circulating narratives and misinformation, and any crossover with other diseases.
- Look beyond high engagement posts to identify narratives within the data.
- Focus not only on the content of conversations but also understanding people's reactions, confusion, questions, and sentiment about health guidance.
- Pay attention to vulnerable populations, stigma, and inequities.
- Perform a visual analysis of memes and images in the media and on social media that are associated with mpox.
- Compare analyses over time to identify changes in narratives and emerging topics.
- Refer to WHO-authored papers (see below) on how to apply your analyses and validate their use.
- Refer to the WHO/UNICEF [How to build an infodemic insights report in 6 steps](#) for advice on how a taxonomy can be used in the context of developing an infodemic insights report, including data analysis methods (see below).
- Remember to continually assess and review your search strategy, including the breadth of search terms, data sources and taxonomy categories.

An important point to note is that a social listening taxonomy is designed to capture conversations to identify narratives. It is not epidemiological and seeks to look beyond the established disease determinants, occurrence and distribution. Defining the categories is useful, but it is important that it is not too prescriptive to exclude the unexpected. As an example, during the COVID-19 pandemic there were rumours that the virus was being transmitted via 5G networks. While COVID-19 is not transmitted this way, capturing these conversations is important and any search strategy needs to be agile enough to incorporate new information that may not be expected.

WHO-authored papers and resources

- WHO/UNICEF How to build an infodemic insights report in 6 steps.
<https://www.who.int/publications/i/item/9789240075658>
- Public Health taxonomy for social listening on respiratory pathogens (includes guidance)
<https://iris.who.int/handle/10665/373534>
- OpenWHO course Infodemic management: Defining a taxonomy for social listening
<https://openwho.org/courses/im-taxonomy>
- Purnat TD, Vacca P, Czerniak C, Ball S, Burzo S, Zecchin T, Wright A, Bezbaruah S, Tanggol F, Dubé È, Labbé F. Infodemic signal detection during the COVID-19 pandemic: development of a methodology for identifying potential information voids in online conversations. JMIR infodemiology. 2021 Jul 28;1(1):e30971. <https://infodemiology.jmir.org/2021/1/e30971>
- Purnat TD, Nguyen T, Ishizumi A, Yau B, White B, Cecchini S, Samuel R, Hess S, Bezbaruah S, Briand S. 2022. Delivering actionable infodemic insights and recommendations for the COVID-19 pandemic response. Weekly Epidemiological Record. 2022 Jul 8; 97 (27), 313 - 324.
<https://apps.who.int/iris/handle/10665/359144>

Taxonomy structure

A taxonomy allows an analyst to monitor and report a narrative summary on millions of conversations, mapped to specific public health categories. The categories outlined here are designed to capture the breadth of conversations and aid in a rapid analysis of their volume and velocity change and enable a narrative trend analysis. The taxonomy spans five topics of conversation:

1. **The cause** – How did the virus emerge and how is it spreading?
2. **The illness** – What do we know about the disease, what are the symptoms and how is it transmitted?
3. **The treatment** – How can it be diagnosed, cured, treated and what is being said about treatments and therapeutics?
4. **The interventions** – What is being done by authorities and institutions?
5. **Conversations about information** – Meta-conversations about guidance, reporting, misinformation and content

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
Category #1: How are people talking about <i>the cause</i>? – Questions, concerns, confusion, information voids, narratives, mis-/disinformation about how the virus emerged and how it is spreading		
The source or cause of the disease or virus		- Includes the original cause, or cause of emergence or reemergence (i.e. mutation). - For example: zoonotic - from animals/ monkeys / small mammals / rodents, smallpox/variola, other. - Possible speculation about bioengineering, health security, laboratory incidents - Defrosting of tundra/environmental sources (smallpox concerns) - Other potentially discussed causes, i.e. caused by vaccines, witchcraft. The analyst can consider whether conversations about the cause or spread is natural or man-made.
Stigma about or by specific populations		- Towards specific population groups: For example: gay, bisexual and other men who have sex with men and sex workers, trans and gender diverse people. - Those from specific geographical locations. - Stigma or discrimination towards vaccinated or unvaccinated people. - Towards those with other visible skin conditions. - These people may be infected with the virus or suspected of spreading the virus, such as healthcare workers. An analyst should be able to identify stigmatizing narratives regardless of whether these are related to a specific population. Stigma may be related to people who may be infected and spreading the virus, as well as perceptions

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
		about those who are not infected, but others may stigmatize as spreading the virus.
Virus spread		Travel: - Urban/rural, geographical spread. - Spread via travel, including cross-border, air travel, animal trade. Community contexts: - Places where transmission is more likely to occur. People living or working in close and crowded places. For example: overcrowded housing, slums, refugee settlements workplaces, farms, school/nursery/childcare, factories, warehouses, markets, health care, aged care, prisons, houses of worship, transportation, sports, festivals, concerts, social gatherings, parties and large events, hunting, meat processing - Unsanitary conditions. The analyst can consider whether the conversations relate to the virus spreading or the host (i.e. a person or animal) moving or travelling. These narratives would focus on conversations about virus spread in terms of communities, community contexts or in relation to other factors, rather than focussing on specific transmission pathways.
Category #2: How are people talking about <i>the illness</i>? – Questions, concerns, confusion, information voids, narratives, mis-/disinformation about what we know about the disease, what the symptoms are and how it is transmitted		
Symptoms	Confirmed	Confirmed: rash, blisters, sores, lesions, headache, fever, swollen lymph nodes, muscle aches, back pain, weakness, proctitis (inflammation inside the rectum), difficulties urinating.
	Other symptoms being discussed	Include other signs and symptoms and confusion with the symptoms of other diseases, for e.g. chickenpox in children. Conversations here can signal confusion that an analyst may need to investigate in detail. It may be useful to identify what is currently unknown to help better understand how people are discussing events.
Means of transmission		This category is titled 'means of transmission' as it includes conversations about confirmed modes of transmission, and any other potential transmission pathways that people are discussing. An analyst can consider monitoring this category over time and splitting it into subtopics if needed. Including person-to-person transmission: For example: face-to-face exposure, through bodily fluids (including sexual contact, sharing needles,

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
		etc.), respiratory droplets, including health care workers without appropriate personal protective equipment (PPE). • Mother to baby, in-utero, breastfeeding (note, although transmission via breastmilk remains unknown, capturing conversations is important). Contaminated materials, organic and inorganic: • For example: contaminated materials such as clothing and bedding or utensils that have been used by a symptomatic person, shared surfaces and tools. Zoonotic/animal-human contact: • Consumption of raw or undercooked animal products including milk and meat or contact with animal faeces. Contact with live or dead animals including handling, scratches/bites, petting, cuddling, hugging, kissing, licking, sharing sleeping areas, and sharing food, as well as handling a dead animal (or parts), skinning, or butchering the animal. Asymptomatic/pre-symptomatic transmission: • This can include the incubation period before showing symptoms, continuing to spread post-symptoms, and complete asymptomatic transmission, or 'carrier'. • Note that while asymptomatic transmission has been reported but not confirmed, capturing conversations is still important. Other: • Conversations about other means of transmission.
Protection from transmission/ disease prevention		Conversations about isolation, hand washing, condom use, marital sex only, breastfeeding, safe preparation of meat products. Physical or physiological barriers or methods that prevent the transmission from occurring. Other conversations on personal measures, vaccination and measures in public settings would be captured in other categories.
Demographics, vulnerability and risk		Risk groups may include: children, infants, pregnant women, persons that are immunocompromised, including people with untreated and advanced HIV disease, close contacts of those who have mpox, health and care workers, sex workers, men who have sex with men, older people.

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
		Analysts should try to capture whether conversations about increased risks are about or by certain groups.
Severity of disease	Severity of disease – personal factors	Disease severity by sociodemographic and health factors: children, infants, pregnant women, persons that are immunocompromised, including people with untreated and advanced HIV disease, older people. Analysts should try to capture whether conversations about disease severity are about or by certain groups.
	Secondary and dual/double infections	Secondary infections: bronchopneumonia, sepsis, encephalitis, secondary bacterial infections of skin, lungs or blood, myocarditis, eye problems. Analysts can consider characterizing people's concerns about an increased risk of secondary infections and impacts (i.e. having a lowered immune system from the first infection).
	Long-term consequences	Includes conversations about the long-term repercussions of illness
Virus comparisons		How mpox is being discussed together in terms of risk and severity in relation to other viruses such as smallpox/variola virus, chickenpox, or avian influenza and COVID-19. How comparisons are drawn in relation to risk perception, the burden on the health system, health delivery and similar.
Virus changes (clades)		For example, discussions about emerging or mutating clades or subclades. Including cause, impact and severity. <i>(Aug 2024: There are two known clades of MPXV: clade I, previously known as the Congo Basin clade, which includes subclades Ia and the recently identified Ib; and clade II, previously called the West African clade; clade II further has two subclades: clade IIa and clade IIb).</i> Includes discussion about how the mpox virus may change and the impact of the changes. Note: narratives about how to outbreak itself is changing would be coded in an alternative category, e.g. the virus spread, demographic or vulnerable population categories
Immunity		For example, various discussions of immunity including smallpox immunity (via vaccination or previous infection), post-infection, herd immunity, 'hybrid immunity', natural immunity, post-vaccination immune response, post-infection immune response (including reinfection).
Category #3: How are people talking about <i>the treatment</i>? – Questions, concerns, confusion, information voids, narratives, mis-/disinformation about how the illness can be diagnosed, prevented or cured and what treatments and therapeutics there are		

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
Treatments	Current treatments	- Home care, rest, hydration, analgesics and antipyretics, sore care (antiseptics, warm baths, salt water rinses) - Antivirals (Tecovirimat)
	Non-proven treatments	Includes therapeutics and treatments under trial but not yet proven (brincidofovir, Tembexa), natural and herbal remedies not proven to be treatments, and others. Treatments used in unapproved ways (i.e. antibiotics for viral infection) and things used as treatments that are not (i.e. methanol consumption, ingesting disinfectant). Include discussions on efficacy, timing, safety and side-effects.
Vaccine (general)		Includes vaccine efficacy, safety, access, prioritisation, stockpiles, recommended scheduling and dose etc. Context-specific analysis is important; for example, narratives of post-exposure prophylaxis and vaccines.
Vaccine side-effects and adverse events		Vaccine side-effects, adverse events after vaccination, vaccine compensation schemes. Includes conversations about both confirmed side-effects and adverse events, and otherwise.
Research and development		Includes conversations about vaccine/treatment nationalism, supply chain challenges, clinical trials, vaccine technology (i.e. live virus), licensing and regulatory approval.
Testing and diagnosis		Including access to, accuracy of, benefit of tests. May include discussions of test type. Discussions about other types of diagnosis.
Category #4: How are people talking about <i>the interventions</i>? – Questions, concerns, confusion, information voids, narratives, mis/disinformation about what is being done by the authorities and institutions. including policies, mandates, recommendations, and advice		
Personal protective equipment (PPE)		For example: gloves, gowns, medical masks and eye protection for health care workers, lab workers; masks, sheets, gowns for patients. Conversations could include discussion of the effectiveness of PPE, its supply, availability and accessibility.
Supportive care		Including health care, health care workers and health care equipment and technology. The analyst can monitor and consider splitting this category if needed.
Personal measures		For example: washing hands, isolation, condom use

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
Inequalities and human rights		For example: perception of and conversations about inequities in service design and measures, delivery and quality of care, access to treatment and vaccines.
Measures in public settings	Travel measures	Note: Even if not currently recommended, people may be discussing them.
	Health and vaccine certificates	Note: Even if not currently recommended, people may be discussing them. This category would also include conversations about mandates regarding testing (i.e. where a negative test is required to access a venue).
	Contact tracing	Case isolation/quarantine duration and close contact definitions.
	Reduction in movement	Isolation and quarantine protocols and recommendations. May include longer-term effects from reduction in movement (i.e. where isolation periods may have an impact on mental health).
	Other measures	Any other discussed intervention measures. Including a ban on animal transportation, isolation of animals or animal culling.
Industry and economic impact		Measures by, or impacting on, industry as well as economic impacts and motivations. This can include impacts on people's livelihoods and work (for example lost wages due to isolation), the cost of disease management, vaccination, and financial motives for measures, for example.
Faith		Includes narratives about faith and religious leaders, faith-based communities in general and beliefs in terms of illness and health.
Category #5: How are people talking about <i>information</i>? – Meta-conversations about guidance, reporting, misinformation and content		
Data, research and evidence		How people are talking about data, research and evidence, including different sources and reliability, etc. Sources may include mass-media, social networks, academic and institutional papers, government policies and datasets, for example. This would include meta-conversations about data, research and evidence in general. Specific conversations relating to data, research and evidence relating to particular topics, would be captured in the other categories.
Health guidance		Reactions, confusion, conversations, feelings and interpretation of guidance and guidelines. Would include

Topic	Subtopic	Notes for the analyst on keyword setup and interpretation
		meta-conversations about conflicting health advice between global jurisdictions. Confusion about specific pieces of conflicting or confusing health guidance should be captured in the category it relates to (i.e. differences in vaccine recommendations in each country, or confusions between treatment advice from different authorities, etc.)
Mis- and disinformation		For example: how people are talking about the existence and management of misinformation, disinformation and malinformation, both offline and online and what their attitudes are towards it. Includes overarching conspiracy narratives of the illness being a scam, a lie or authorities or other groups planning population control, for example. It may be useful to know if particular groups or organizations are being characterized as sources of misinformation. Other misinformation and disinformation concerns. This category captures meta-conversations. For example, how people are talking about the management of misinformation by social media platforms, rather than specific pieces of misinformation which would be included in the taxonomy topics (e.g. vaccines are implanting microchips, etc.).

Reference on the mpox 2024 outbreak:

- <https://www.who.int/publications/m/item/multi-country-outbreak-of-mpox--external-situation-report-35--12-august-2024>
- <https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern>
- <https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON522>
- <https://www.who.int/emergencies/disease-outbreak-news/item/2024-DON525>
- <https://www.who.int/news-room/questions-and-answers/item/monkeypox>