

## An anti-infodemic virtual center for the Americas

Ian Brooks,<sup>1</sup> Marcelo D'Agostino,<sup>2</sup> Myrna Marti,<sup>2</sup> Kate McDowell,<sup>1</sup> Felipe Mejia,<sup>2</sup> Miguel Betancourt-Cravioto,<sup>2</sup> Lisa Gatzke,<sup>3</sup> Elaine Hicks,<sup>4</sup> Rebecca Kyser,<sup>1</sup> Kevin Leicht,<sup>5</sup> Eliane Pereira dos Santos,<sup>2</sup> Jessica Jia-Wen Saw,<sup>3</sup> Ailin Tomio,<sup>6</sup> and Sebastian Garcia Saiso<sup>2</sup>

**Suggested citation** Brooks I, D'Agostino M, Marti M, McDowell K, Mejia F, Betancourt-Cravioto M, et al. An anti-infodemic virtual center for the Americas. *Rev Panam Salud Publica*. 2023;47:e5. <https://doi.org/10.26633/RPSP.2023.5>

### ABSTRACT

The Pan American Health Organization/ World Health Organization (PAHO/WHO) Anti-Infodemic Virtual Center for the Americas (AIVCA) is a project led by the Department of Evidence and Intelligence for Action in Health, PAHO and the Center for Health Informatics, PAHO/WHO Collaborating Center on Information Systems for Health, at the University of Illinois, with the participation of PAHO staff and consultants across the region. Its goal is to develop a set of tools—pairing AI with human judgment—to help ministries of health and related health institutions respond to infodemics. Public health officials will learn about emerging threats detected by the center and get recommendations on how to respond. The virtual center is structured with three parallel teams: detection, evidence, and response. The detection team will employ a mixture of advanced search queries, machine learning, and other AI techniques to sift through more than 800 million new public social media posts per day to identify emerging infodemic threats in both English and Spanish. The evidence team will use the EasySearch federated search engine backed by AI, PAHO's knowledge management team, and the Librarian Reserve Corps to identify the most relevant authoritative sources. The response team will use a design approach to communicate recommended response strategies based on behavioural science, storytelling, and information design approaches.

### Keywords

Public Health Informatics; social media; artificial intelligence; COVID-19; communication; Americas.

The COVID-19 pandemic and its response have been accompanied by a massive infodemic, the overabundance of information—some accurate and some not—that occurs during an epidemic, when rapid decisions are necessary. This infodemic has made it harder for governments, institutions, and individuals to make accurate decisions that affect the public health emergency management process and the health of the population (1–3).

An infodemic also involves a large increase in the volume of information on a given topic, which can increase exponentially when an incident such as the COVID-19 pandemic occurs (4,5). In this situation, accurate scientific and technical information

that by itself can be overwhelming and difficult to collect, process, and aggregate is shared, moving through journalistic outlets, social media, etc. Through sharing, accurate information becomes mixed with rumours, manipulated data, fake expertise, incorrect information, and false and biased news, hindering the recipient's ability to process and judge it all. Furthermore, access to false or incorrect data produces significant distortions in predictive models, affecting health system planning and decision-making (6).

Therefore, a holistic and collaborative response framework is needed to properly address the infodemic accompanying COVID-19 or any other public health emergency and to support

<sup>1</sup> Center for Health Informatics, School of Information Sciences, University of Illinois Urbana-Champaign. ✉ Ian Brooks, [ianb@illinois.edu](mailto:ianb@illinois.edu)

<sup>2</sup> Pan American Health Organization, Washington, D.C. United States of America

<sup>3</sup> National Center for Supercomputing Applications, University of Illinois Urbana-Champaign, Urbana, United States of America

<sup>4</sup> The Rudolph Matas Library of the Health Sciences, Tulane University Health Sciences Center, New Orleans, United States of America

<sup>5</sup> Department of Sociology, University of Illinois Urbana-Champaign, Urbana, United States of America

<sup>6</sup> Mind, Behavior and Development Unit, World Bank, Washington, D.C. United States of America

public health policy, strategy or program aiming to advance the right to health and improve health outcomes.

The purpose of this article is to present an overview of the PAHO/WHO Anti-Infodemic Virtual Center for the Americas (AIVCA), an international project led by the Department of Evidence and Intelligence for Action in Health, PAHO, and the Center for Health Informatics, PAHO/WHO Collaborating Center on Information Systems for Health, at the University of Illinois (designation, USA-450), with the participation of PAHO staff and consultants across the Region of the Americas. The goal of the project, which took place between May and December 2021, is to develop a set of tools that pair Artificial Intelligence (AI) with human judgment to help public health decision-makers and their risk communication teams respond to infodemics. The team of senior-level health informaticists, medical librarians, information scientists, sociologists, human behavior specialists, a storytelling expert, and others, were presented with a challenge by PAHO – is it possible to detect emerging infodemic threats to public health, determine the truth behind these threats and make recommendations to ministries of health on how to respond, all within a one week period?

### CONCEPTUAL DESIGN OF THE CENTER

The center is designed with two underlying principles:

1. The primary audience for the center will be ministries of health or associated organizations and their surveillance and risk communication teams in the Americas.

2. The design and implementation must be as scalable as possible, so that the need for human intervention is specified and minimal to ensure efficiency.

The virtual center is structured by three fundamental steps to infodemic threat management: 1. Detection, 2. Evidence, and 3. Response.

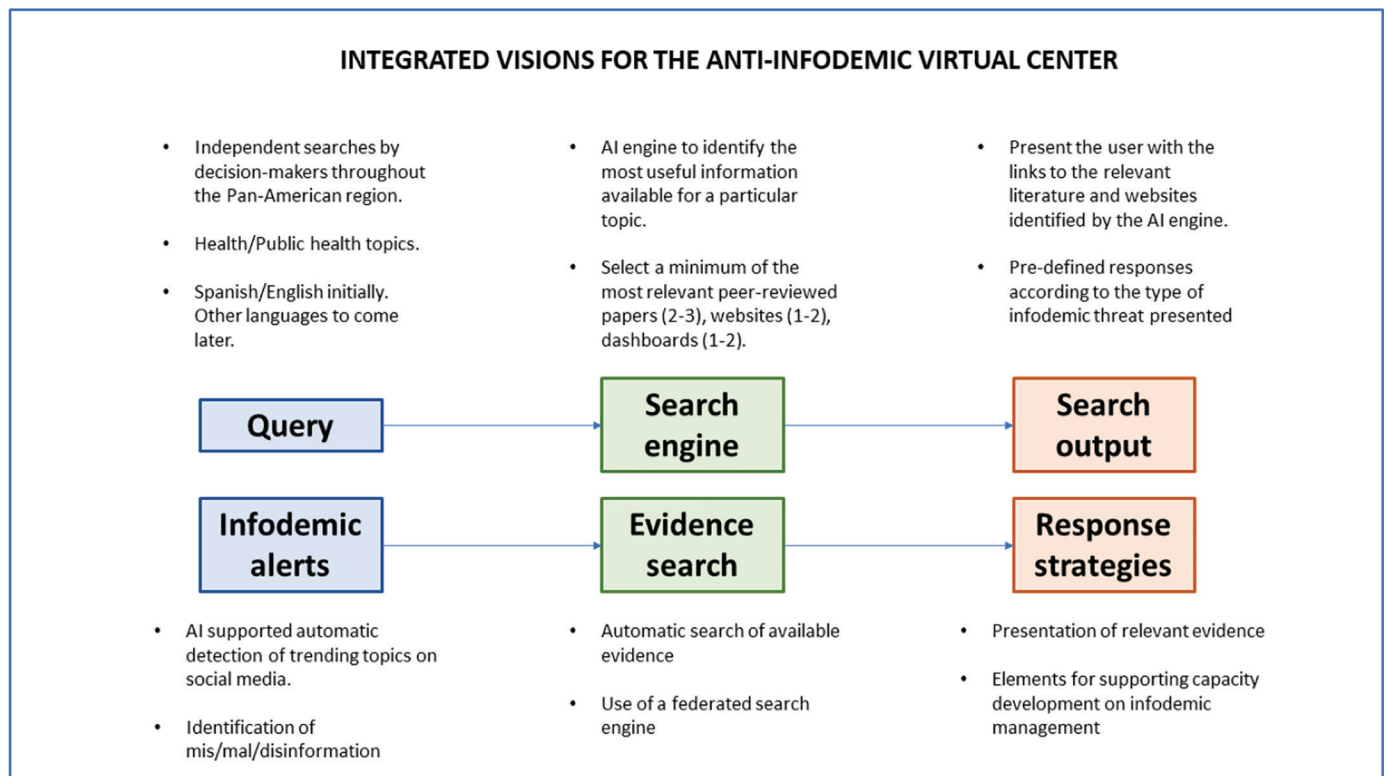
Different ministries have different needs and capabilities, so the center is designed to be flexible in its use and the level of support it provides. It is anticipated that there will be two main use cases both related to infodemic threats – stories being spread that have the potential to cause harm to public health:

1. A threat has been identified and a ministry of health or other public health organization wants help in evaluating and responding or
2. A ministry of health or other public health organization wants help detecting emerging threats.

Therefore, the AIVCA consists of two parallel services, figure 1:

1. A search interface. Users will input search terms related to topics relevant for public health decision-making. The system will then carry-out a wide search of relevant peer-reviewed literature, grey literature, and websites of authoritative organizations and institutions. From the results of the search, an AI engine will select (3-5) relevant papers, documents and websites that will then be presented to the user. The user will also be offered the choice of seeing response recommendations. If the threat has already been examined by the evidence

FIGURE 1. Anti-infodemic virtual center of the Americas – conceptual design



Source: This figure was prepared by the authors based on the word described in this paper.

and response teams, then a previously created response may be presented, otherwise general recommendations will be related to classification of different types of threat.<sup>1</sup>

2. A misinformation alert. This alert presents the output of a series of AI-supported processes for the detection of dis/mis/mal information topics with potential public health impact for countries in the region. The search for topics will be carried out weekly in social media outlets and other relevant sources. After a topic is selected, the teams will present to the users: a) the (3-5), most relevant papers, documents and websites, and b) a series of behavioral sciences-based recommendations on how to better develop an anticipatory response.

Initially, the AIVCA will work in Spanish and English, with the aim to expand the system to work in the rest of PAHO's official languages (specifically Portuguese and French). Further, at this time AI, taken in its broad definition to encompass machine learning, neural networks, and deep learning etc., cannot be the complete solution. However, due to the scope of the problem, AI has to be part of the solution and indeed, recent work has demonstrated methods detecting fake news and misinformation (8,9).

## IMPLEMENTATION OF THE CENTER DESIGN

Based on its conceptual design, the AIVCA includes three sequential components corresponding to the three fundamental steps of infodemic threat management: 1. Detection, 2. Evidence, and 3. Response. A team of experts was constituted to address each of the steps.

### COMPONENT 1: DETECTION

The objective of the detection component is to rapidly and reliably identify new and emerging mis/dis/mal-information threats to public health in time for ministries of health and other organizations to prepare and implement mitigation strategies.

For the purposes of the proof-of-concept implementation the team developed three interconnected algorithms focusing on text-based (English and Spanish) social media:

1. A machine learning approach to monitor high-level topics of discussion on social media for further investigation, primarily Twitter, but also including Reddit, blogs, YouTube comments, etc.
2. An AI approach that can detect linguistic signatures of probable misinformation posts.
3. Automated monitoring of social media recommendations algorithms on sites such as YouTube, TikTok, etc. along with other infodemiology sites such as WHO EARS, Google Fact Check, Junkipedia and COVID19Misinfo.

### ALGORITHM 1. TOPIC MODELLING

The primary data source for obtaining public social media posts in the target languages is Twitter due to its volume. Posts to Reddit, blogs, YouTube comments are also included. Detailed

<sup>1</sup> We acknowledge that the need for transparency, reproducibility, and the related concept of "explainability" (7) will require details about the AI engine that are beyond the scope of the present overview of this project.

### Examples of Twitter Search Strings

((NOT Abbott) ((COVID OR Vaccine\* OR mask\* OR Variant OR Delta OR Vax\* OR Vaccinated OR Iota OR Kappa OR Mu OR "Covid vaccine"~10) NEAR/20 (Cure\* OR medicine OR truth OR immunity OR lies OR treatment\* OR mask\* OR treats\* OR natural\*)) NOT Airlines NOT sororities NOT fraternities NOT "greek life"

((NOT Abbott) ((corona OR coronavirus OR pandemia OR coronavac OR covid\* OR Delta OR Vax\*OR Iota OR Kappa OR Mu OR "vacuna"~10) NEAR/20 (cura OR trata\* OR medic\* OR inmunidad OR mascar\* OR verdad OR en serio\* OR menti\* OR verd\* OR natural\*)) NOT aerolínea NOT sororities NOT comunidad NOT "vida grieg"

Preprocessing steps included removing links and special characters, converting emojis and emoticons to meaningful text, tokenization and stemming.

search queries in both English and Spanish were developed to reduce the ~800 million new posts per day (10) that the team has access to, to ~50 000 – 350 000 relevant posts per day to download for further analysis with algorithms 1 and 2.

Topic modeling produces an overview of the main themes behind the week's new posts. This unsupervised machine learning technique analyses large amounts of text data by assigning categories based on the topic or theme of each individual paragraph. It can infer patterns and group similar statements without the need to predefine subject tags or provide training data. Latent Dirichlet Allocation (LDA) was used to perform the topic modelling (11).

### ALGORITHM 2. NEURAL NETWORKS

A central element to the development of the AIVCA is exploring the role that Artificial Intelligence and Machine Learning can play in misinformation detection. The overall objectives of this endeavour are to:

1. Examine the overall quality of data sources and their fitness for machine learning and AI applications using standard API downloading procedures.
2. Engage with external facts-checking organizations whose principal purpose is to provide facts-checking for journalists (for this project Poynter (12) and Meedan (13)), access their compiled data, and construct AI models that identify social media posts as misinformation, focusing on grammatical structure and word usage.
3. Expand modelling efforts beyond these curated training datasets to examine un-curated posts on Twitter and Facebook.
4. Create a process that will identify misinformation as belonging to one of the specific categories: a. misapplication of science, b. conspiracy theories or c. pseudo-science (14).

Ultimately, continually refined AI/Machine Learning (ML) tools will help to identify significant health-related information threats that undermine trust in public health infrastructure. The main advantage of AI/ML tools, once refined, is their potential

to identify emerging threats as they happen. These tools will also be able to assess the volume of the threat identified in terms of the number, voracity, and spread of social media posts related to a specific misinformation topic.

### ALGORITHM 3. SOCIAL MEDIA TRENDING AND INFODEMIOLOGY PEER SITE MONITORING

Sharing social media platforms prevent search or download of the vast majority of posts because of privacy settings, so the ability to detect misinformation threats from these sources is limited. It is possible, however, to take advantage of the platforms' own algorithmic reporting to know what is trending. For instance, both TikTok and YouTube, through analysis of their algorithms autoplay of similar content, can provide data on trending misinformation. Data-processing workflows to download the identified videos, extract the audio tracks, and convert them into text for analysis are being developed.

In addition to these misinformation extraction methods, monitoring methods will follow what other infodemiology (15) and fact checking sites are detecting and extract posts to include in our text analysis stream. These sites include WHO EARS, Junkipedia, Google Fact Check and COVID19Misinfo.

### DETECTION OUTPUT

These three algorithms will each identify a number of threats every week. The choice of which threat to focus on for the evidence and response teams will be determined by the three teams taking into account multiple factors including the prevalence of the threat and the speed of its dissemination, the judged risk to public health, and the ability of ministries of health to mount a meaningful response. Behavioral science will be central to this decision.

### COMPONENT 2: EVIDENCE

The objective of the Evidence component is to develop a process to identify sound and authoritative evidence to refute or confirm trending viral mis/dis/mal information. Evidence will appear in both parallel services, via the search interface or the misinformation alert (figure 1).

The Evidence component is based on information sciences and centered on determining the best tools and sources available to quickly define and retrieve the most relevant sources as 3-5 search results (scientific articles, websites, dashboards, etc.) to support or reject contested information, if such sources exist. It is important to use a systematic review and a published literature search method, in order to communicate transparency and reproducibility required for state levels of authority.

Members of the recently formed Librarian Reserve Corps (16) will refine searches and provide expert guidance to select the results to be included in the output provided via the two parallel services. The interface will support searching and the misinformation alert will provide information on an AIVCA detected threat.

To be useful, the search engine needs to access literature from a wide variety of disciplines, since debunking misinformation threats may need evidence from many fields outside health. For example, disproving a claim that the vaccine causes patients to

become magnetic requires a knowledge of physics as well as of vaccine composition, and direct evidence may not exist since the misinformation in this case is scientifically absurd. Human judgment will be necessary. The Evidence team will use, a database of databases (or federated search engine) to conduct an advanced search. The federated search engine (FSE) will gather information from the following resources:

- Virtual Health Library (including LILACS)
- WHO Global COVID-19 database
- PUBMED/MEDLINE
- Others that meet the AIVCA evidence criteria, are publicly available, and allow appropriate data harvesting for transparency of search process.

An extensive FSE, EasySearch (17) has already been designed by the University of Illinois and customized to meet project needs in partnership with the University of Illinois Library and Center for Health Informatics.

Challenges identified in setting up the Evidence component include:

- Difficulty proving that there is no evidence. The mere search of bibliographic databases cannot confirm the absence of evidence to disprove misinformation. Underlying issues include that authoritative evidence is not produced fast enough to disprove trending viral topics, evidence that is considered authoritative may change rapidly during an emergency, and confabulatory conspiracies may not have any counter-evidence.
- Absence of evidence. Lack of sources in recorded and/or published literature (peer-reviewed, in pre-print databases and grey literature) is not an indication that trending viral information is false.
- Speed. Carrying out quick but thorough searches in bibliographic databases storing journal articles, grey literature, monographs, reference books, etc. is difficult.

To face those challenges, it is necessary to develop a machine-processed literature review or a process that will be at least as fast as the detection algorithm. Speed alone will not solve all evidence issues, however, and human judgment will be necessary as the evidence moves forward to response.

### COMPONENT 3: RESPONSE

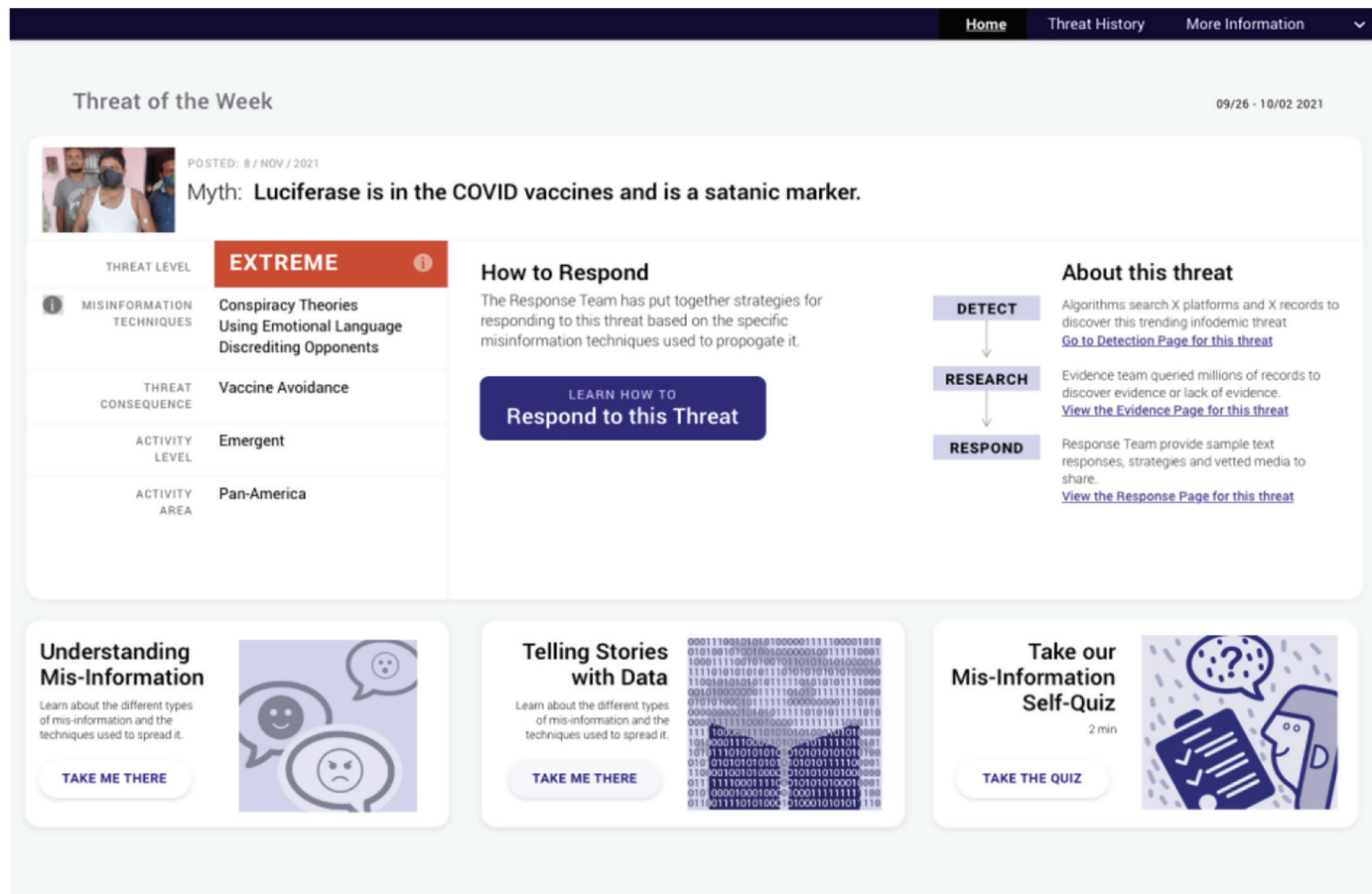
The aim of the Response component is to present AIVCA users with the output from the detection algorithms plus the selection of appropriate available evidence as a misinformation alert. This alert is for decision-makers in ministries of health and associated institutions as the end user. The response products (figures 2 and 3) are based on behavioral and communication sciences, storytelling techniques, information design, and user interface and experience design approaches and best practices.

The response product includes:

1. Alert. Warning of emerging mis/dis/mal information topics that may represent a risk to people's health.
2. Explanation. Reasons for considering a probable public health risk posed by an infodemic topic. To assess the risk the following variables are used: Data (subvariable: Magnitude



FIGURE 2. Threat of the Week and general information



Source: This figure was prepared by the authors based on the word described in this paper.

- or size of the problem) and Public Health Impact (sub variables: Significance, Vulnerability and Susceptibility).
- Evidence. Description of what scientific evidence is available -or not- for or against the specific topic under analysis. Evidence is framed using storytelling and behavioral techniques that acknowledge the emotion (e.g., fear, anger) found in social media posts, and presented as succinct “What we know” snippets. Each snippet could be expanded with a dropdown feature to show associated shareable content, vetted by the evidence team and curated by the response team.
  - Strategies. What to do to counteract possible negative impacts of this type of misinformation topic. The response prioritizes educating underlying concepts in mis/dis/mal information over prescribing actions applicable only to a single threat.

Whether searching for a topic (search interface) or viewing the latest threat (misinformation alert), users will be able to visualize the response product as a series of virtual cards that present the topic, the estimated risk level, the evidence and arguments for or against, and the misinformation classification and behavioral response strategies, (figure 2). Virtual cards and other shareable content are presented with a purple “share” icon, which cues decision-makers to copy or share content with a single mouse click or tap on mobile device (figure 3). This feature is designed to minimize the barrier for decision-makers

to take in content and share with their communities. The figures that follow give an example based on recent misinformation related to the term “luciferase.”

Additionally, the Response component includes content to help decision-makers detect potential mis/dis/mal information independently from the algorithms of the evidence team. The design of this content was informed by recent literature (18) on the behavioral aspects of misinformation, which differentiates the act of debunking (retroactively reducing reliance on misinformation) from prebunking (pre-emptively reducing the persuasiveness of misinformation). Prebunking can be achieved via interventions based on the inoculation theory (18,19). This theory suggests that in order to psychologically inoculate a person, two actions are required: i) warning them of an impending threat to their beliefs, and ii) presenting a refutation of the arguments brought up by the threat and highlighting the fallacies and techniques used. Even though debunking strategies have shown effectiveness in the past (20,21), they have shown to have limitations over continued influence of misinformation. Therefore, prebunking strategies have gained attention as another effective way to counter misinformation and its rapid viralization.

Material is created to help increase awareness of emerging concepts in mis/dis/mal information, as well as how to respond to each type of information. This includes (1) a one minute self-test on recognizing mis/dis/mal information, (figure 4) and (2)

FIGURE 3. Proposed shareable virtual card for the Response product

✕

Posted: 09/26/2021

## RESPONSE STRATEGIES

**Misinformation:**  
The COVID-19 vaccine contains the substance Luciferase and Luciferase is a satanic marker.

### What we know

None of the COVID-19 vaccines contain Luciferase. >

The word "luciferase" comes from a Latin root—lucifer, in Latin, means "light-bearing"—and has also been used by poets to refer to Venus, the morning star. >

The similarity between the enzyme name "luciferase" and the angel/devil name "Lucifer" from old stories has sparked fear in some people of Covid-19 vaccines. >

The enzyme luciferase is a natural substance, and it's a normal part of the wonder of lightning bugs. <

**Content to Share**

**Fireflies are a natural source of the enzyme called Luciferase.**  
—Anti-Infodemic Resiliency Center

The enzyme luciferase is a natural substance, and it's a normal part of the wonder of lightning bugs. The name might sound bad, but that's a coincidence that should not cause fear. Scientists sometimes name discoveries, in this case in 1888, with words that sound different or scary a century and more later, but there's no need for fear. All language changes over time, and this is one example. Knowing how people are scared of the word now, it's unlikely that scientists would use the same language today, but they didn't know how we might interpret the word. And there's nothing bad about this enzyme, it's just part of the wonder of creation that some creatures can glow naturally!

**Trusted Sources to Share**

<https://www.tmc.edu/news/2020/07/fireflies-help-...>

**Fireflies help kindle new tests and treatments for COVID-19**  
Jul 21, 2020 — Located in the heart of Houston, Texas Medical Center campus is home to leaders in research, medicine, and innovation in healthcare.

<https://www.scientificamerican.com/article/how-and-...>

**How and why do fireflies light up? - Scientific American**  
Sep 5, 2009 — When oxygen combines with calcium, adenosine triphosphate (ATP) and the chemical luciferin in the presence of luciferase, a bioluminescent ...

[https://en.wikipedia.org/wiki/Firefly\\_luciferase](https://en.wikipedia.org/wiki/Firefly_luciferase)

**Firefly luciferase - Wikipedia**  
Firefly luciferase is the light-emitting enzyme responsible for the bioluminescence of fireflies and click beetles. The enzyme catalyzes the oxidation of ...  
Organism: Photinus pyralis UniProt: P08659  
PDB: 1LQI  
Mechanism of reaction · Structure · Spectral differences in... · Homology

<https://www.businessinsider.com/science/video>

**What makes a firefly glow - Business Insider**  
Jun 30, 2021 — The firefly is probably the Earth's most famous bioluminescent species, with over 2000 individual species. Here's how they make their light.

### General Information

**Understanding Mis-Information**

Learn about the different types of mis-information and the techniques used to

TAKE ME THERE

**Telling Stories with Data**

Learn about the different types of mis-information and the techniques used to

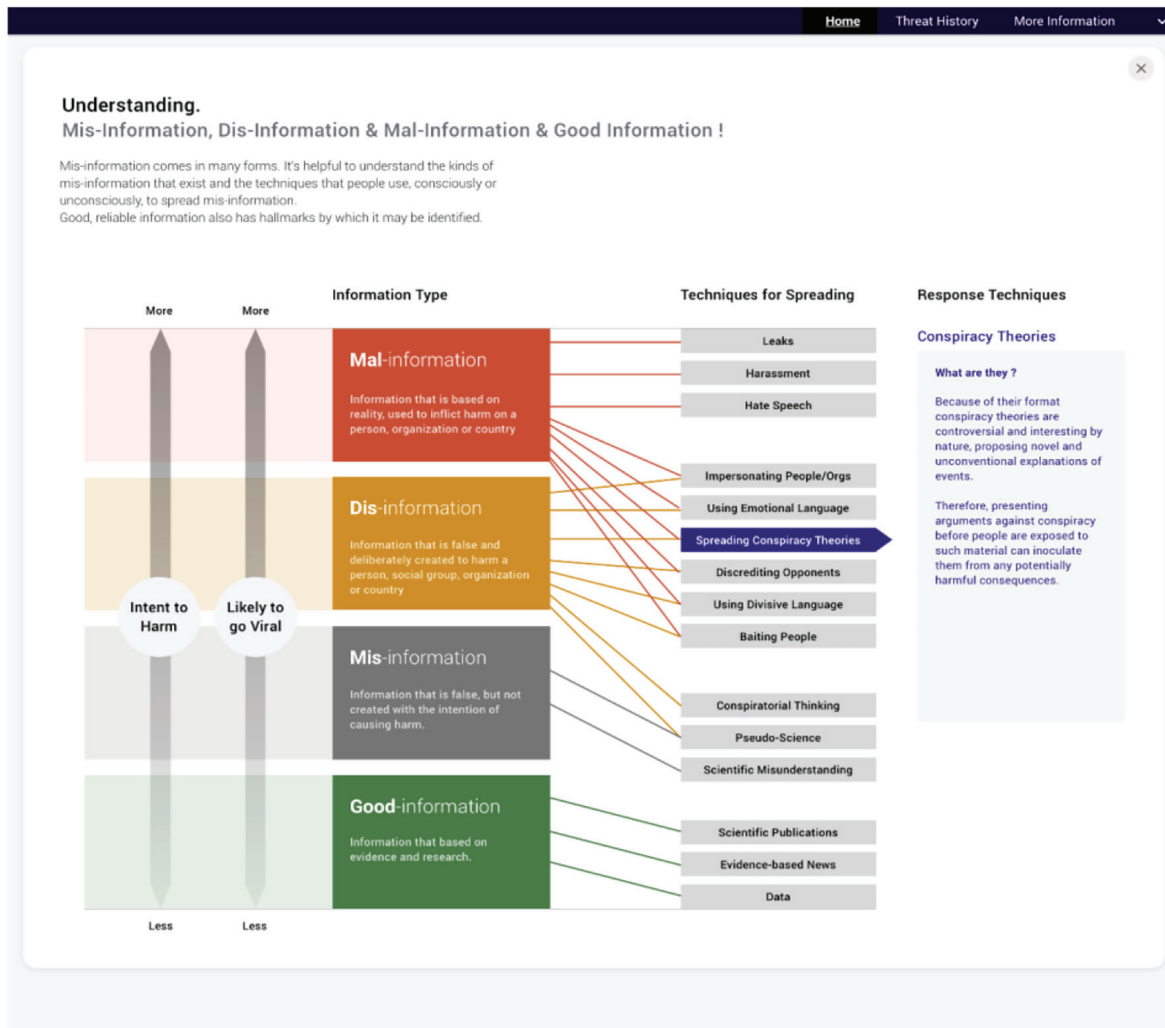
TAKE ME THERE

Source: This figure was prepared by the authors based on the word described in this paper.

FIGURE 4. One minute self-test on recognizing mis/dis/mal informatio

Source: This figure was prepared by the authors based on the word described in this paper.

FIGURE 5. Infographic for understanding mis/dis/mal information and associated response techniques



Source: This figure was prepared by the authors based on the word described in this paper.

an infographic for understanding mis/dis/mal information and the response techniques associated with each type of information (figure 5). The self-test questions and misinformation techniques were extracted from a novel five-minute browser game called *Go Viral!*, which was developed and evaluated by behavioral specialists(18). *Go Viral!* has shown to improve people’s attitudinal confidence on their ability to detect COVID-19’s misinformation and reduce self-reported willingness to share the misinformation with others.

**STORYTELLING AS A CENTRAL ELEMENT FOR INFODEMIC MANAGEMENT**

An effective Response component will provide users with the right information through the right channels in an easily understood manner. The AIVCA uses storytelling ways of thinking for infodemic management, which is a novel way of seeing how story sharing and shaping happens, with an understanding of what storytellers are doing when they communicate (22).

The AIVCA collaboration and design process revealed that misinformation detection and response cannot be fully automated at this time. What is needed is a systematic hybrid human-machine way to help leaders in public health get a handle on what has gone wrong when misinformation is spreading quickly.

Stories contain both informational and emotional content. When people are sharing/posting/retweeting/etc., the teller-audience-story dynamic means that people are sharing ways of making sense, whether accurate or misinformed. Storytelling explains aspects of current crises in truth that have challenged prior AI and human interception efforts. Serious consideration of storytelling may be a missing piece in such efforts to build systems—both human and computational—that thwart misinformation such as COVID-19 infodemic issues and more.

Efforts to stop misinformation in all its forms must empower health leaders as storytellers. Reaching misinformed audiences requires story-ready information that health leaders can connect



with the audiences they know. Accurate information that can be easily retold will prepare them to deal with infodemic threats as directly and simply as possible. This includes empathy and care for the informational and emotional experiences of audiences. Addressing audiences where they are, as people who are misinformed in particular ways or fearfully grasping specific untrue stories, will allow health leaders' recommendations to be heard. It is not always possible to know why large audiences keep sharing and retelling false stories. However, it is possible to prepare health leaders to focus on the agency of the audience in choosing what information to share, as inoculation or as counter-narratives to misinformation.

## DISCUSSION

The proof-of-concept design of the AIVCA allows for the development and testing of a complex, scalable, multilingual human-with-algorithm process for detecting new and emerging misinformation threats using a combination of statistical, ML, and AI techniques. To date, the components have been designed and piloted, with successful demonstration of the ability to detect and respond to threats.

There is still work to be done, however, mainly in the integration and automation of the many component sub-algorithms, further testing of the neural network-based linguistic signature detector, and development of better methods to analyse image and video social media. The progress made enables confidence that it is possible to detect misinformation threats from social media with minimal human intervention and quickly enough to provide public health organizations time to respond.

For the Evidence component, human-based work helped to create a sound search methodology based on international data search standards, and next steps will automatize the process as much as possible by refining the selection algorithm of documents from the scientific literature and websites, creating automated search processes and protocols to translate detected threats into usable evidence and engage subject headings, and completing the integration of LILACS with EasySearch. Next steps will include defining and streamlining the needs for human judgement in the Evidence component. It is also necessary to implement a tool to gather and analyze user feedback on results provided for system improvements.

Next steps in the link between the Evidence and Response components require adapting a data storytelling framework as a tool for determining what evidence is needed to support or rebuke an infodemic topic or when the available evidence is not helpful. This will include:

- Testing and refining the newly developed Misinformation Detection Rubric with input from end users

- Piloting and refining the risk assessment matrix for classifying threat levels with user input
- Developing default inoculation technique recommendations based on the category of misinformation: conspiracy theory, pseudoscience, or misunderstood science
- Refining scoring algorithms for ranking threats

## CONCLUSION

The AIVCA is the result of progress in the theory and characterization of a rather new social event, the infodemic. It has also made it possible to integrate scientific knowledge from diverse areas to create an effective response to the potentially negative effects of the infodemic, such as data science for the detection team, information science for the evidence team and behavioral sciences for the response team.

Despite its complexity, the infodemic and what its management implies, the AIVCA will present public health decision makers with a simple but effective management solution expressed in a Response (misinformation alert) with a friendly design where mis/dis/mal information trends are shown and characterized, confronted against available evidence, and offered mechanisms to counteract such trends with the aim of protecting the population's health.

**Author Contributions.** MD, and MM had the original concept for the virtual center. All authors contributed to the design and planning of the center. IB, FM, and KM wrote the paper. All authors reviewed and approved the final version.

**Acknowledgements.** The authors would like to thank Prof Bill Mischo for assistance with EasySearch; Prof Peg Burnett, Jessica Callaway, Brad Long and other members of the Librarian Reserve Corps; and the following students who provided valuable input during the implementation of this proof-of-concept: (by alphabetical order) Sreenithi Balasubramanian, Nikita Bangale, Anushri Bhagwath, Rashmi Chhabria, Shruti Deekshitula, Kaivan Gandhi, Jerry Guo, Poorva Joshi, Sarthak Joshi, Asmita Khode, Steven Li, Wanpeng Liu, Zahra Malwi, Atulya Mannava, Vaibhavi Mehta, Faiz Mohammed Faiz, Yunji Nam, Nithin Narla, Nikhil Obili, Shubhangi Ranjan, Saurav Rathod, Rohan Salvi, Dhruvesh Shah, Maitri Shah, Ankita Sharma, Sushanth Sreenivasa Babu, Shriya Srikanth, Qianqian Wang, Shen Yu

**Conflicts of interest.** None declared

**Disclaimer.** Authors hold sole responsibility for the views expressed in the manuscript, which may not necessarily reflect the opinion or policy of the *Revista Panamericana de Salud Pública / Pan American Journal of Public Health* and/or those of the Pan American Health Organization.

## REFERENCES

1. Tangcharoensathien V, Calleja N, Nguyen T, Purnat T, D'Agostino M, Garcia-Saiso S, et al. Framework for managing the COVID-19 infodemic: methods and results of an online, crowdsourced WHO technical consultation. *J Med Internet Res*. 2020;22(6):e19659.
2. D'Agostino M, Medina Mejía F, Martí M, Novillo-Ortiz D, Hazrum F, de Cosío FG. Infoxicación en salud. La sobrecarga de información sobre salud en la web y el riesgo de que lo importante se haga invisible. *Rev Panam Salud Publica*. 2017;41:e115.
3. Dias P. From 'infoxication' to 'infosaturation': a theoretical overview of the cognitive and social effects of digital immersion. *Ámbitos Revista Internacional de Comunicación*. 2014 (24). URL: <https://revistascientificas.us.es/index.php/Ambitos/article/view/9979>

- [accessed 2022-8-30] Author please Review this reference and cite accordally
4. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M. Assessing the risks of 'infodemics' in response to COVID-19 epidemics. *Nat Hum Behav.* 2020;4(12):1285–93.
  5. Pan American Health Organization. COVID-19 Factsheet: Understanding the Infodemic and Misinformation in the fight against COVID-19 [Internet]. Washington D.C.: PAHO; [April 2020]. Available from: <https://iris.paho.org/handle/10665.2/52052>, <https://iris.paho.org/handle/10665.2/52053> (Spanish), <https://iris.paho.org/handle/10665.2/52054> (Portuguese), <https://www.paho.org/sites/default/files/2020-10/podcast-covid-08-en.mp3> (Podcast)
  6. García-Saisó S, Marti M, Brooks I, Curioso W, González D, Malek V, et al. The COVID-19 Infodemic. *Rev Panam Salud Publica.* 2021;45:e56. <https://doi.org/10.26633/RPSP.2021.56>.
  7. Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable AI: a review of machine learning interpretability methods. *Entropy.* 2020;23(1):18. 2020 Dec 25;23(1):18. doi: 10.3390/e23010018.
  8. Szczepański M, Pawlicki M, Kozik R, Choraś M. New explainability method for BERT-based model in fake news detection. *Sci Rep.* 2021;11(1):23705. doi: 10.1038/s41598-021-03100-6.
  9. Kolluri N, Liu Y, Murthy D. COVID-19 Misinformation Detection: Machine-Learned Solutions to the Infodemic. *JMIR Infodemiology.* 2022;2(2):e38756.
  10. Internet Live Stats. Twitter Usage Statistics. [2022-May-17] Available from: <https://www.internetlivestats.com/twitter-statistics/>
  11. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res.* 2003;3:993–1022.
  12. Poynter Poynter. [Internet]. [cited 2022 May 17]. Available from: <https://www.poynter.org/>
  13. Meedan Meedan. [Internet]. [cited 2022 May 17]. Available from: <https://medan.com/>
  14. Kyser R, McDowell K, Betancourt-Cravioto M, D'Agostino M, Gatzke L, Hicks E, et al. Categorization of infodemic threats. *Forthcoming.* 2023.
  15. Eysenbach G. Infodemiology: the epidemiology of (mis)information. *Am J Med.* 2002;113(9):763–5.
  16. Callaway J. The Librarian Reserve Corps: An Emergency Response. *Med Ref Serv Q.* 2021;40(1):90–102. Available from: <https://doi.org/10.1080/02763869.2021.1873627>.
  17. Mischo WH, Schlembach MC, Bishoff J, German EM. User search activities within an academic library gateway: implications for web-scale discovery systems. In: Popp MP, Dallis D, editors. *Planning and implementing resource discovery tools in academic libraries.* Hershey, PA: IGI Global; 2012. Available from: <https://doi.org/10.4018/978-1-4666-1821-3.ch010>
  18. Basol M, Roozenbeek J, Berriche M, Uenal F, McClanahan WP, Linden S van der. Towards psychological herd immunity: cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data Soc.* 2021;8(1):20539517211013868.
  19. Compton J. Inoculation theory. In: Dillard JP, Shen L, editors. *The SAGE handbook of persuasion: developments in theory and practice.* New York: SAGE Publications Inc.; 2012. Available from: [https://sk.sagepub.com/reference/hdbk\\_persuasion2ed/n14.xml](https://sk.sagepub.com/reference/hdbk_persuasion2ed/n14.xml)
  20. Ecker UKH, Antonio LM. Can you believe it? An investigation into the impact of retraction source credibility on the continued influence effect. *Mem Cogn.* 2021;49(4):631–44.
  21. Guillory JJ, Geraci L. Correcting erroneous inferences in memory: the role of source credibility. *J Appl Res Mem Cogn.* 2013;2(4): 201–9.
  22. McDowell K. Storytelling wisdom: Story, information, and DIKW. *J Assoc Inf Sci Technol.* 2021;72(10):1223–33.

Manuscript submitted on 23 May. Revised version accepted for publication on 19 September 2022.

## Un centro virtual contra la infodemia para la Región de las Américas

### RESUMEN

El centro virtual contra la infodemia para la Región de las Américas de la Organización Panamericana de la Salud/Organización Mundial de la Salud (OPS/OMS) es un proyecto liderado por el Departamento de Evidencia e Inteligencia para la Acción en la Salud de la OPS y el Center for Health Informatics de la Universidad de Illinois, centro colaborador de la OPS/OMS en sistemas de información para la salud, con la participación de personal y consultores de la OPS en toda la Región. Su objetivo es crear un conjunto de herramientas que combinen inteligencia artificial (IA) y los criterios humanos para apoyar a los ministerios de salud y las instituciones relacionadas con la salud en la respuesta a la infodemia. Los funcionarios de salud pública recibirán formación sobre las amenazas emergentes detectadas por el centro y recomendaciones sobre cómo abordarlas. El centro virtual está estructurado en tres equipos paralelos: detección, evidencia y respuesta. El equipo de detección empleará una combinación de consultas mediante búsqueda avanzada, aprendizaje automático y otras técnicas de IA para evaluar más de 800 millones de publicaciones nuevas en las redes sociales al día con el fin de detectar amenazas emergentes en el ámbito de la infodemia tanto en inglés como en español. El equipo de evidencia hará uso del motor de búsqueda federado EasySearch y, con el apoyo de la IA, el equipo de gestión del conocimiento de la OPS y la red Librarian Reserve Corps, determinará cuáles son las fuentes autorizadas más pertinentes. El equipo de respuesta utilizará un enfoque vinculado al diseño para difundir las estrategias recomendadas sobre la base de las ciencias del comportamiento, la narración de historias y el diseño de la información.

### Palabras clave

Informática en salud pública; medios de comunicación sociales; inteligencia artificial; COVID-19; comunicación; Américas.

---

## Um centro virtual anti-infodemia para as Américas

### RESUMO

O Centro Virtual Anti-Infodemia para as Américas (AIVCA, na sigla em inglês) da Organização Pan-Americana da Saúde/Organização Mundial da Saúde (OPAS/OMS) é um projeto liderado pelo Departamento de Evidência e Inteligência para a Ação em Saúde da OPAS e pelo Centro de Informática em Saúde da Universidade de Illinois, EUA (Centro Colaborador da OPAS/OMS para Sistemas de Informação para a Saúde), com a participação de funcionários e consultores da OPAS de toda a região. Seu objetivo é desenvolver um conjunto de ferramentas — combinando a inteligência artificial (IA) com o discernimento humano — para ajudar os ministérios e instituições de saúde a responder às infodemias. As autoridades de saúde pública aprenderão sobre as ameaças emergentes detectadas pelo centro e obterão recomendações sobre como responder. O centro virtual está estruturado com três equipes paralelas: detecção, evidência e resposta. A equipe de detecção utilizará consultas de pesquisa avançada, *machine learning* (aprendizagem de máquina) e outras técnicas de IA para filtrar mais de 800 milhões de novas postagens públicas nas redes sociais por dia, a fim de identificar ameaças infodêmicas emergentes em inglês e espanhol. A equipe de evidência usará o mecanismo de busca federada EasySearch, com apoio de IA, da equipe de gestão de conhecimento da OPAS e do Librarian Reserve Corps (LRC), para identificar as fontes abalizadas mais relevantes. A equipe de resposta usará uma abordagem de design para comunicar estratégias de resposta recomendadas com base em abordagens de ciência comportamental, narração de histórias e design da informação.

**Palavras-chave** Informática em saúde pública; mídias sociais; inteligência artificial; COVID-19; comunicação; América.

---