Digital Communication

Navigating the COVID-19 pandemic:
Artificial intelligence, natural language generation, and the COVID-19 Tracking Project

SECTION EDITOR

Nicole Bezuidenhout
njbezuidenhout@gmail.com

Matthew Rector
Associate Vice President
Narrativa
matthew.rector@narrativa.com

Sofía Sánchez González
Publicity and Marketing
Narrativa
sofia.sanchez@narrativa.com

The past two years have confronted humanity with a variety of unprecedented challenges due to the far-reaching COVID-19 pandemic. In response to this tragedy have come monumental new advances in science and technology – namely the vaccines developed to prevent COVID-19 and curb the spread of the virus, as well as the detection methods used to identify it. However, an area that the public might not naturally associate with scientific and technological progress related to COVID-19 is artificial intelligence. Artificial intelligence (AI) is a subject matter with which most are already tangentially familiar. It is used every day when interacting with our mobile devices, or when deciding what film or TV programme to stream. Even marketing messaging is now produced using AI! But what one may not know is that a subfield of AI, called natural language generation (NLG), has played a pivotal role in our response to the COVID-19 pandemic. NLG has been essential in assessing massive amounts of COVID-19-related data to, among other things, provide accurate and easy-to-understand information to multiple stakeholders around the world.

Simply put, NLG is a computer software process that collects and transforms raw data into written natural human language. Though this technology is widely used by many businesses and organisations in a variety of sectors, it has proven particularly useful in the life sciences industry during these unprecedented times. To this point, the automation of clinical study reports, patient safety narratives, and electronic common technical documents, better known as eCTDs, have enabled medical writers to perform faster and improve efficiencies. Pharmaceutical companies have also substantially reduced costs and accelerated regulatory submissions, bringing much needed new medical developments to the market quicker. With these and other useful applications in mind, the COVID-19 Tracking Project was created to further contribute to the COVID-19 response.

About the project

The COVID-19 Tracking Project is an international initiative that aims to facilitate the transformation of data about the virus into essential knowledge using NLG. It is a collaborative effort between Los Angeles-based Narrativa, one of the global pioneers in the NLG space, and several international partners including: the Spanish international news agency Agencia EFE; the Spanish public corporation for public radio and television services Corporación de Radio y Televisión Española (RTVE); the online news outlet Infobae; the information provider Applied XLabs; the location intelligence platform Carto; and the design experience and visual storytelling company DesignIt.

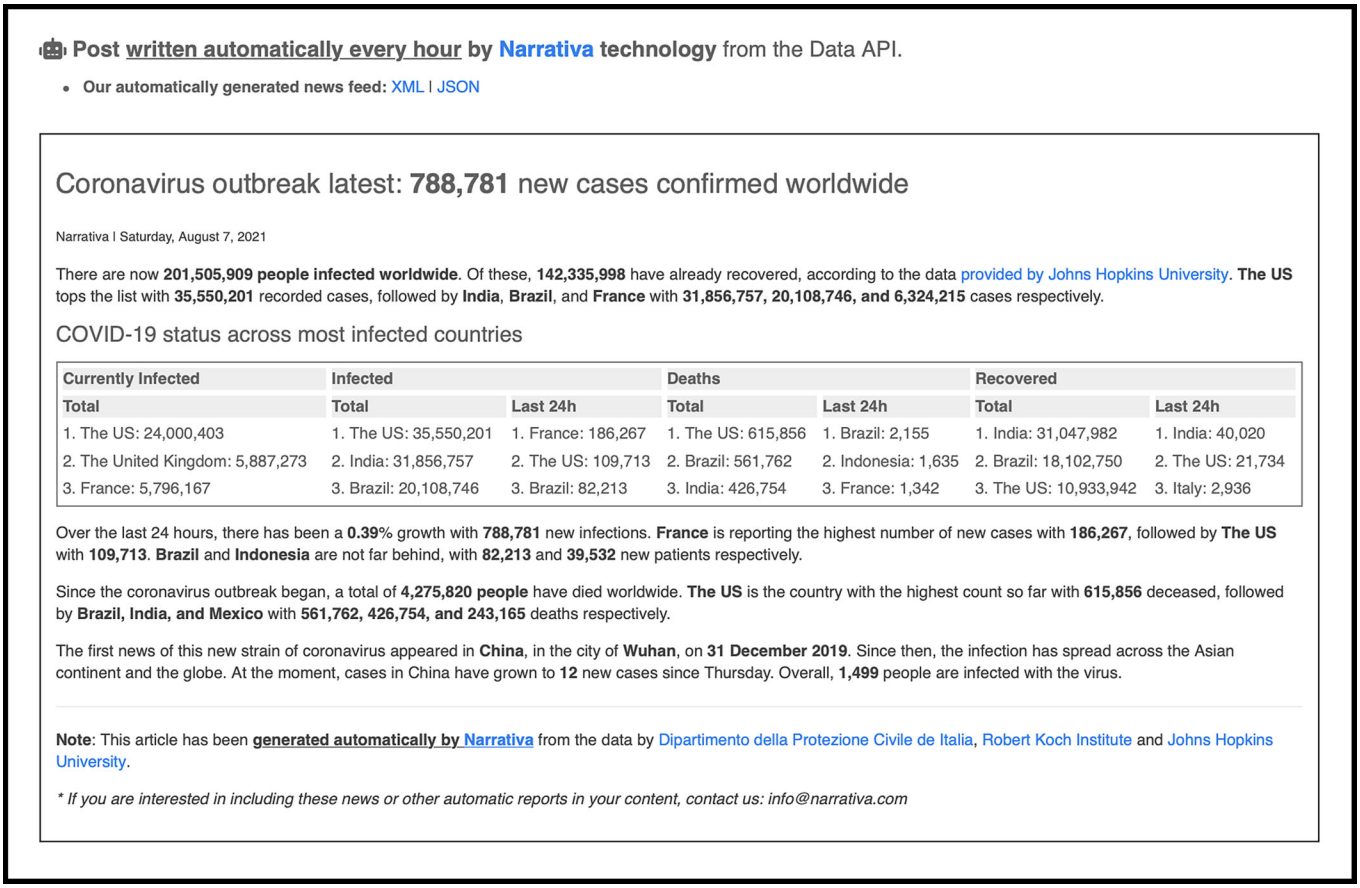
> Such easy access to accurate information about COVID-19 has been especially useful given the volume of inaccurate and false information that has been circulating among online communities.

"Throughout the pandemic, it seemed as if everyone was drowning in data and did not fully understand what the data surrounding COVID-19 was actually saying. We sought to empower institutions, the media, and the public through natural language generation by giving them the critical knowledge they needed. It's our mission to better the world with technology and we saw an opportunity to help," says Narrativa President Jennifer Bittinger.

More about the tool

Data are first independently collected and verified for accuracy by various health authorities themselves, including Germany's Robert Koch Institute,[1] the Dipartimento della Protezione Civile in Italy,[2] and Johns Hopkins University in the United States.[3] The tool then aggregates the verified data from these sources and structures it into a legible and interpretable format. Daily tests are performed to validate that the data extraction process is correct and in working order. For instance, if any data source changes its format in a way that negatively impacts the tool's software, the tests fail, and the appropriate parties are notified so they can manually intervene and resolve the issue quickly. The software records every update in a database accessible via an application programming interface, commonly known as an API. Finally, and most importantly, AI and NLG convert the data into clear and intelligible text and a chart (Figure 1). The entire process happens automatically and 24/7. Other types of graphic content that can be generated include images and banners which, like the text and chart outputs, can be quickly disseminated and understood by multiple audiences, including the media and public.

Such easy access to accurate information about COVID-19 has been especially useful given the volume of inaccurate and false information that has been circulating among online communities. Moreover, the tool allows one to check the current COVID-19 statistics for any country on Earth, or, more granularly, for various regions, states, autonomous communities, and provinces. Aside from offering a general summary of the data, the information is represented visually through an interactive map



Post written automatically every hour by Narrativa technology from the Data API.

- **Our automatically generated news feed:** XML | JSON

Coronavirus outbreak latest: **788,781** new cases confirmed worldwide

Narrativa | Saturday, August 7, 2021

There are now **201,505,909 people infected worldwide**. Of these, **142,335,998** have already recovered, according to the data provided by Johns Hopkins University. **The US** tops the list with **35,550,201** recorded cases, followed by **India**, **Brazil**, and **France** with **31,856,757, 20,108,746, and 6,324,215** cases respectively.

COVID-19 status across most infected countries

Currently Infected	Infected		Deaths		Recovered	
Total	Total	Last 24h	Total	Last 24h	Total	Last 24h
1. The US: 24,000,403	1. The US: 35,550,201	1. France: 186,267	1. The US: 615,856	1. Brazil: 2,155	1. India: 31,047,982	1. India: 40,020
2. The United Kingdom: 5,887,273	2. India: 31,856,757	2. The US: 109,713	2. Brazil: 561,762	2. Indonesia: 1,635	2. Brazil: 18,102,750	2. The US: 21,734
3. France: 5,796,167	3. Brazil: 20,108,746	3. Brazil: 82,213	3. India: 426,754	3. France: 1,342	3. The US: 10,933,942	3. Italy: 2,936

Over the last 24 hours, there has been a **0.39%** growth with **788,781** new infections. **France** is reporting the highest number of new cases with **186,267**, followed by **The US** with **109,713**. **Brazil** and **Indonesia** are not far behind, with **82,213** and **39,532** new patients respectively.

Since the coronavirus outbreak began, a total of **4,275,820 people** have died worldwide. **The US** is the country with the highest count so far with **615,856** deceased, followed by **Brazil, India, and Mexico** with **561,762, 426,754, and 243,165** deaths respectively.

The first news of this new strain of coronavirus appeared in **China**, in the city of **Wuhan**, on **31 December 2019**. Since then, the infection has spread across the Asian continent and the globe. At the moment, cases in China have grown to **12** new cases since Thursday. Overall, **1,499** people are infected with the virus.

Note: This article has been **generated automatically by Narrativa** from the data by Dipartimento della Protezione Civile de Italia, Robert Koch Institute and Johns Hopkins University.

** If you are interested in including these news or other automatic reports in your content, contact us: info@narrativa.com*

Figure 1. Screenshot of the COVID-19 Tracking Project tool.
The tool contains easy-to-comprehend text and a simple chart derived from the mass amounts of data related to the virus.

and shows the worst hit countries alongside the number of daily positive cases being documented. Another helpful feature of the project is the information alert system, which notifies users of the latest news on COVID-19. For example, if there is a notable increase or decrease in the number of cases within a region of interest, an alert is automatically generated. This feature is particularly helpful because accurate information and fixed content again reaches the public directly, eliminating the inevitable alteration of information as it passes from person to person and is taken as fact.

Narrativa continues to offer reporting that is updated at least every hour, as the insights provided by the COVID-19 Tracking Project have been vital in helping people make more informed decisions about COVID-19. "The efforts of the COVID-19 Tracking Project have left humanity better prepared when a global pandemic happens again," echoes Bittinger. To reach a vast amount of people, the simple yet comprehensive tool is currently directly available to the public in English, Spanish, and Italian. This endeavour is not-for-profit, so any person or institution can freely use the information provided by the COVID-19 Tracking Project, which has become a large and expansive repository of data, reports, graphs, and images associated with the evolution of the virus. To interact with the tool from the COVID-19 Tracking Project, please visit covid19tracking.narrativa.com.

Conflicts of interest

The authors are associated with Narrativa, which has created and sells as a service the technology discussed in this article.

Data availability statement

For inquiries about data and other supplemental information, please contact info@narrativa.com.

References

1. GitHub. Robert Koch-Institut. 2021 [cited 2021 Aug 7]. Available from: https://github.com/robert-koch-institut.
2. GitHub. Presidenza del Consiglio dei Ministri – Dipartimento della Protezione Civile. 2021 [cited 2021 Aug 7]. Available from: https://github.com/pcm-dpc.
3. ArcGIS Online. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). 2021 [cited 2021 Aug 7]. Available from: https://www.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6.

Author information

Matthew Rector develops and leads cross-functional multimedia and data analytics initiatives. He currently serves as an Associate Vice President at Narrativa and recently worked with the UN's World Food Programme on two global public awareness campaigns.

Sofía Sánchez González heads Narrativa's publicity and marketing efforts. She is a public relations and marketing strategist who builds and maintains alliances with the media and journalists.

In QUEST of better science communication

Nicole Bezuidenhout
njbezuidenhout@gmail.com

Reaching an audience has never been easier. Still, your message could be lost on your audience if not tailored to them. This is especially true when communicating important scientific topics to the public that could significantly impact their daily lives, like climate change or vaccines. Indeed, we observed the influence public health messaging can have on decision-making that drives population behaviour change throughout the COVID-19 pandemic. As such, science communicators may feel inspired to reassess how and to whom they convey information. But where to start?

Recognising the gap for effective dialogue between science and the wider public, a European Horizon 2020 project, QUality and Effectiveness in Science and Technology (QUEST), worked to develop tools, recommendations, and guidelines for communicators in the fields of journalism, social media, and museums. Focusing on vaccines, artificial intelligence, and climate change the QUEST team engaged with communication professionals across Europe to identify specific challenges and opportunities for skills development. The *QUEST communication checklist for scientists: communicating science to the public* represents one of several resources in the QUEST toolkit available for download at https://questproject.eu/.[1] Also of interest, is their *Checklist for science communicators on social media*.[2] With such handy tools and guidelines available, we can all look forward to more high-quality science communication.

quest

Checklist for scientists: communicating science to the public

Jacopo Pasotti, Ilda Mannino, Alessandra Fornetti
Venice International University

1. Are you allocating enough ***time*** for preparation and improving your communication skills?
2. Do you know who your ***audience*** is?
3. Have you identified the key ***messages*** to deliver and kept the focus on them?
4. Are you ***framing the message*** so that it sparks curiosity and is compelling to the public?
5. Are you ***linking to current facts*** or events?
6. Are you communicating ***something that you also care about***?
7. Are you ***telling stories*** or just delivering a list of facts or numbers?
8. Are you using ***a simple explanation***, can it be understood by the public?
9. Are you using ***short sentences***?
10. Have you strategically planned ways to open ***a dialogue and interact*** with your audience?
11. Are you carefully thinking about how to ***keep your delivery or writing lively*** and monitoring the public's reaction?
12. Have you set out strategies to ***deal with scepticism or distrust***?
13. Have you ***practiced*** your communication with non-experts?
14. Are you in touch with your ***communication (or press) officers***?

The above checklist developed within the H2020 QUEST project supports scientists in delivering their message and fine-tuning their communication skills to the public. It is the result of several focus groups and interviews with both scientists working in different fields and trainers experienced in science communication. The checklist also draws from the literature review of published studies on scientists' perception and from available guidelines on science communication. Some of the points in the checklist apply in specific contexts and not in others: scientists are encouraged to go through the list and choose those elements that apply to their circumstances.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 824634

Acknowledgements

A special thanks to the QUEST team for permitting the reprint of the *Checklist for scientists: communicating science to the public* poster in our December 2021 issue.

Conflicts of interest

The author declares no conflicts of interest.

References

1. The QUEST project. 2021 [cited 2021 Oct 12]. Available from: https://questproject.eu/toolkits/.
2. The QUEST project. 2021 [cited 2021 Oct 12]. Available from: https://questproject.eu/checklist-for-science-communicators-on-social-media/.