

BBC NEWS LABS

GoURMET and Multilingual Journalism at News Labs

Sevi Sariisik, Lei He, Anna Blaziak
18 November 2021 - Gourmet User Day

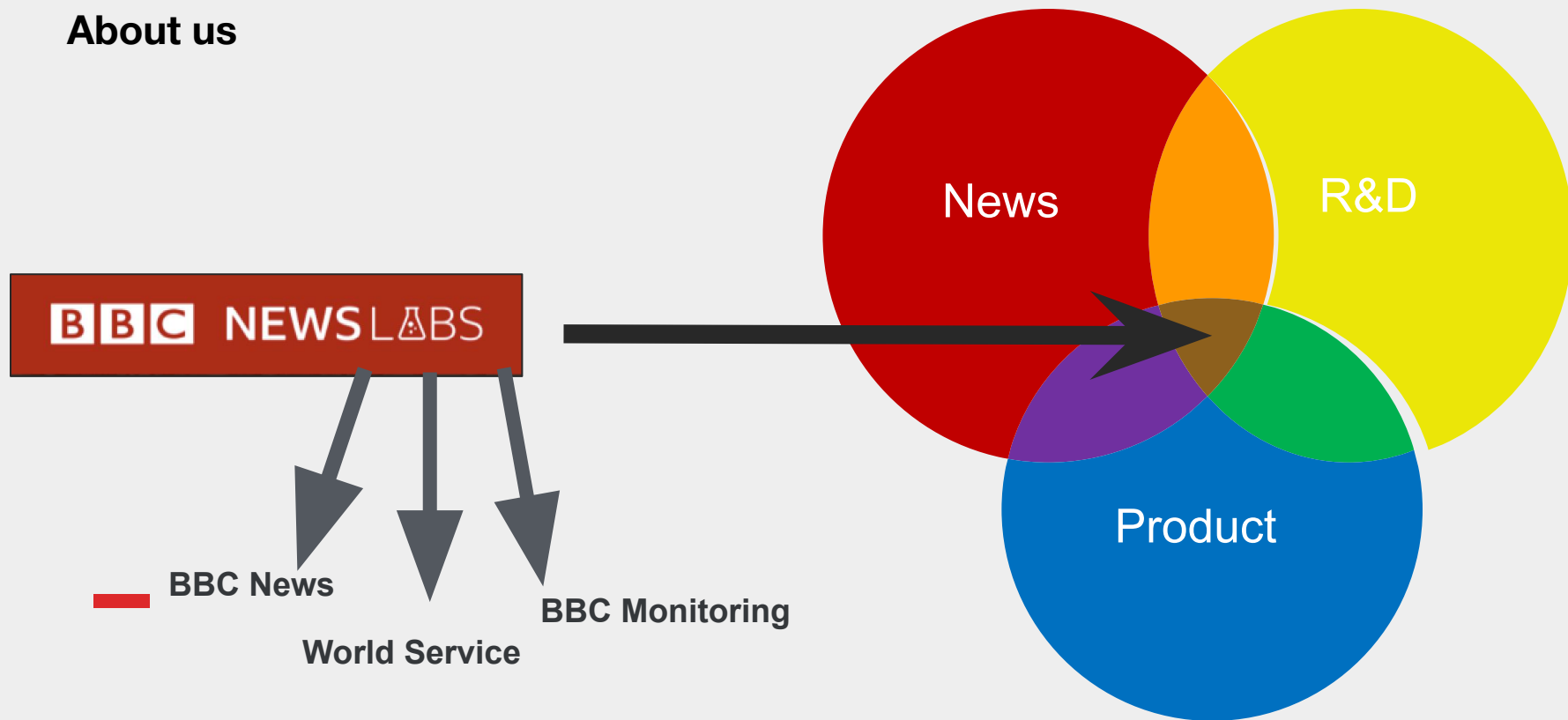
What is multilingual journalism?

How can we work better in a truly multilingual environment?

**Where would multilingualism generate
most value for audiences?**



About us



NL Multilingual Journalism Team



Lei He



Anna Blaziak



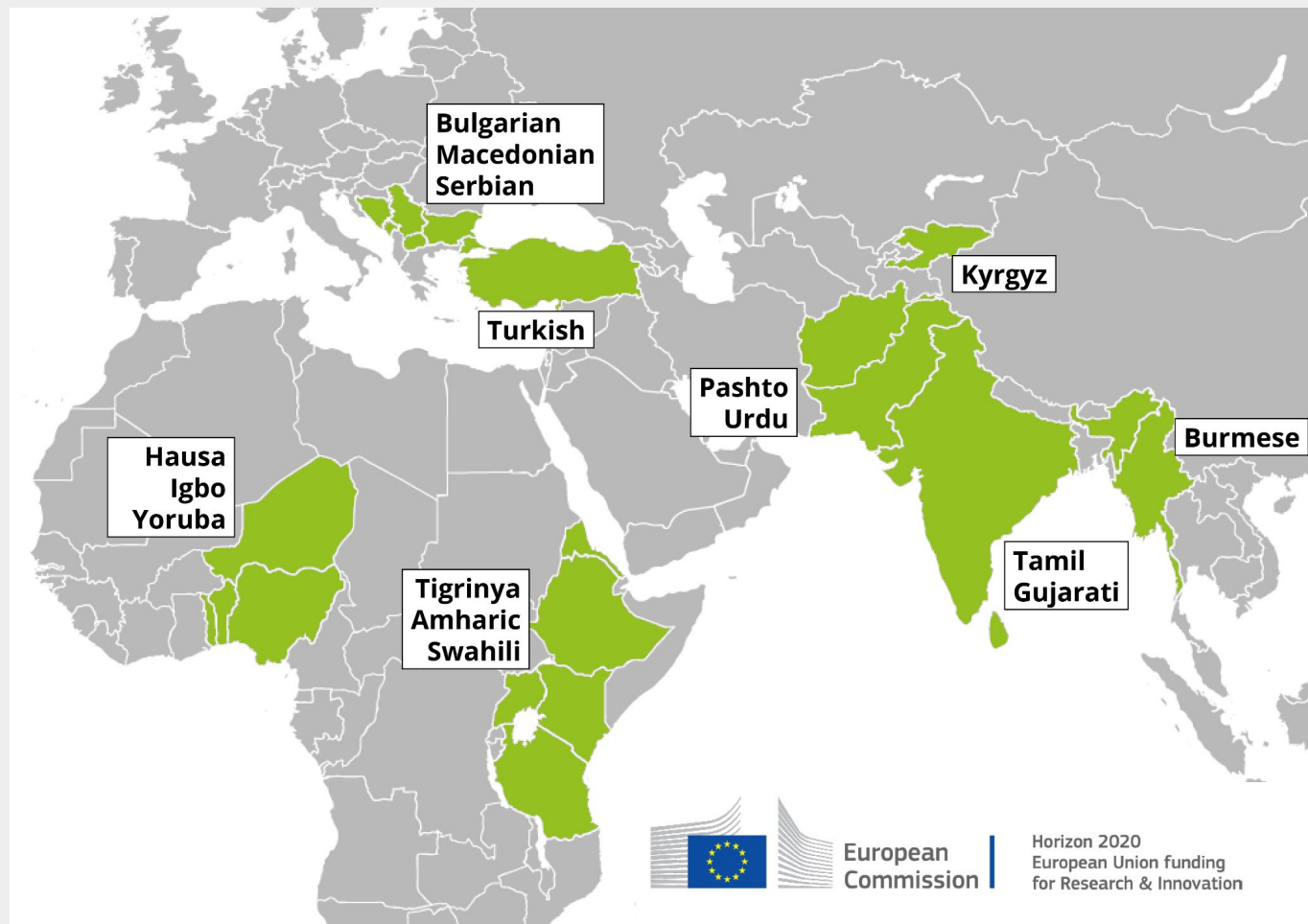
Sevi Sariisik

Scope

2018-2022

Academia-Media
Partnership

16 languages from
BBC & DW
portfolio



European
Commission

Horizon 2020
European Union funding
for Research & Innovation

Languages in scope:

Amharic
Bulgarian
Burmese
Gujarati

Hausa
Igbo
Kyrgyz
Macedonian

Pashto
Serbian
Swahili
Tamil

Tigrinya
Turkish
Urdu
Yoruba



THE UNIVERSITY
of EDINBURGH



Universitat d'Alacant
Universidad de Alicante



UNIVERSITEIT VAN AMSTERDAM



European
Commission

Horizon 2020
European Union funding
for Research & Innovation

Benefits

Close cooperation between

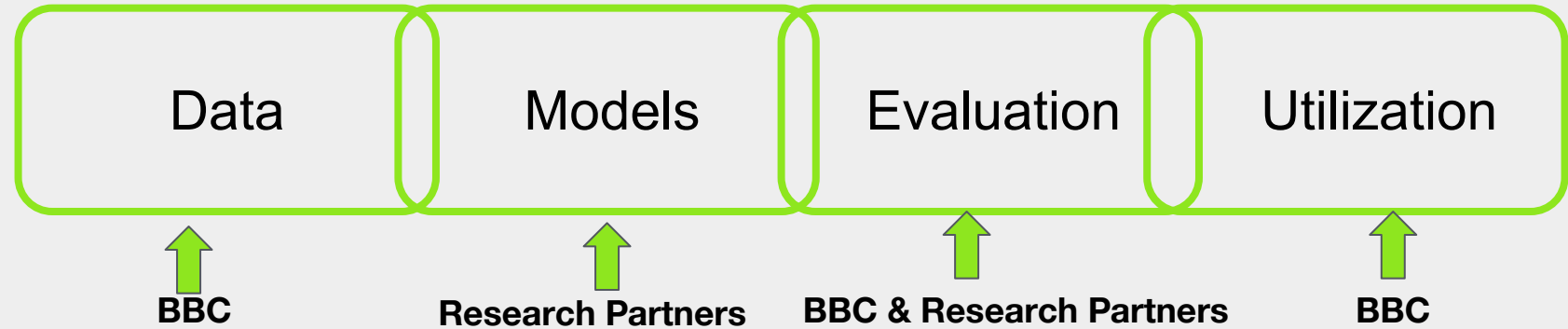
Researchers
Innovation teams
Editorial teams

Cross-discipline exposure & insight

Custom MT models



Workflow



Objectives/Key areas

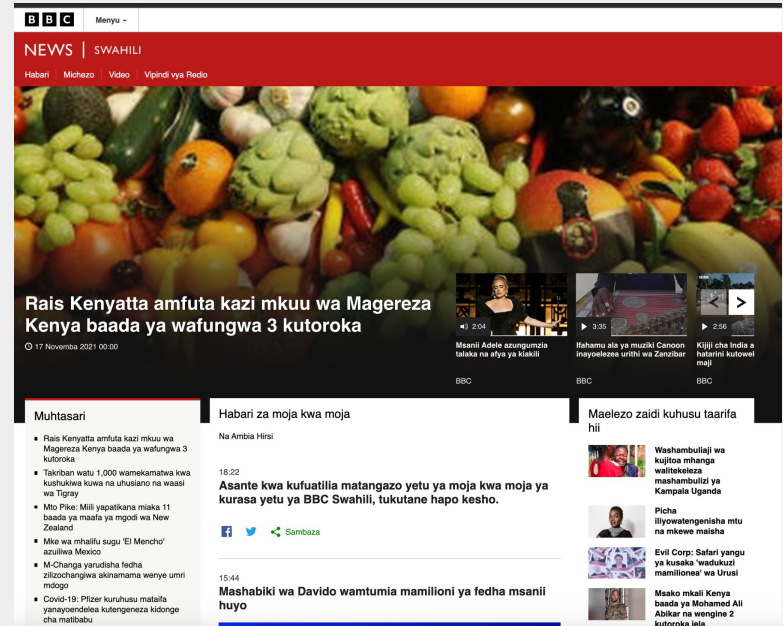
Monitoring		Live Pages Tr.
Content creation		Frank
Domain specific (Health)		Multilingual GST

Live Pages Translation



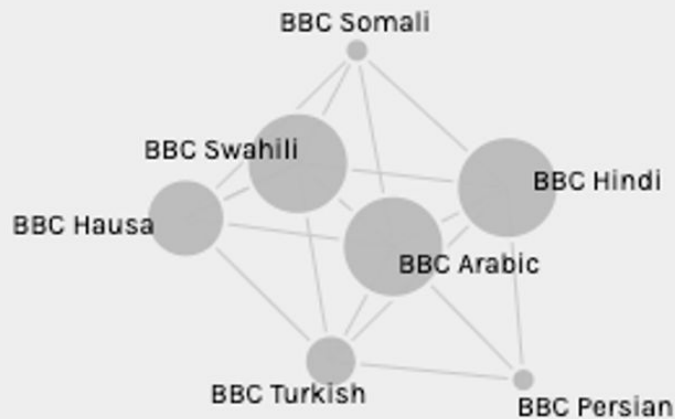
What are Live Pages?

- For popular, fast moving, multi-aspect events/themes, e.g. Coronavirus, Climate Change Conference, Black Lives Matter, etc.
- Written in small snippets, less formal style, fast moving
- Can showcase localized content, that core outlets did not cover



Opportunities

- BBC produces content in **43** languages and most World Service teams run Live Pages
- Teams are interested in content from other teams and they're using **machine translation** already
- **Duplication** in Live Page creation for World Service journalists
- Good original content gets **trapped** in the original language, if other journalists don't know it exists.



Live Pages Translation

Hypothesis: Machine translation is good enough so that journalists want to use it to monitor, create or update live pages on a daily basis.

What: A monitoring tool that allows BBC World Service journalists to easily monitor Live Pages from all World Service teams and to reuse stories.

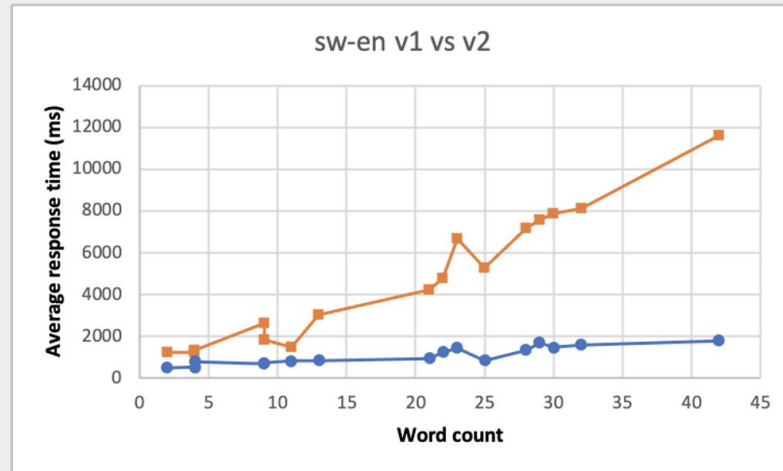
How: A rapid prototype that uses Google Translate with the option to integrate GoURMET models

Trial during U.S. Election and key insights

- **General translation quality:**
 - Machine translations regarded as **high enough quality** for the task required: scanning for content to further research and re-use. Occasional errors reported.
 - Machine translation is not good for non-English language pairs, e.g. from Arabic to Urdu
- **Stories re-used:** Some stories taken and repurposed from other live pages. Arabic, English, Hausa and Azeri all mentioned.
- **Frequent use:** Some users self-report using prototype several times a week.
- **Non-live page content:** Several users interested in seeing other BBC content translated in the feed such as CPS articles. Several would like copyright, video and social media embeds pulled through to feed.

GoURMET model integration

- Processing time for a typical Live Page post (of 150 words) is 6 - 7 seconds.
- We have about 100 posts across all languages per day.
- To ensure acceptable page load time, we would need to pre-translate rather than use on-demand batch translation.
- Collaborated with Juan Antonio from the University of Alicante to reduce the processing time of Swahili models



GoURMET and News Labs

Core digital offer:
Text Articles (150-800 w)

App. 400 articles p/d



BBC NEWS | HAUSA

Labaran Duniya Wasanni Nishadi Cikakkun Rahotanni Bidiyo Shiry-shirye na Musamman Shiry-shiryen rediyo


Dubban daliban Jihar Kano na cikin rashin tabbas saboda bashin gwamnati

Wannan yanayi da daliban suka tsinci kansu na da alaƙa da gazawar gwamnatin jihar na biyan kuɗaɗen da hukumar shirya jarrabawar kammala sakandare ta Najeriya wato NECO ke bin gwamnatin
Sa'o'i 3 da suka wuce


Daliban da ke cikin kunci saboda rashin zuwa makaranta a Kano
Naira miliyan 300 na marayu sun yi batan dabo a Kotun Musulunci a Kano




Real Madrid za ta jinkirta sayen Mbappe, Man Utd za ta rabu da Henderson
Sa'o'i 3 da suka wuce




Masu fyade na cin karensu babu babbaka a Najeriya - Amnesty
Sa'o'i 3 da suka wuce




Birnin da Musulmai suke jagoranci a Amurka
Sa'o'i 3 da suka wuce




Na tabbata dan uwana ya samu kyakkyawar makoma – Kanen Janar Zirkushu
Sa'o'i 3 da suka wuce




'Boko Haram na rike da ikon wasu yankunan jihar Neja'
16 Nuwamba 2021



Yan bindigar da ke kai hari Sokoto za su dan-dana kudarsu - Buhari
16 Nuwamba 2021



Dalilin da ya sa na gina wa Fulani makiyaya makaranta - Sheikh Gumi
16 Nuwamba 2021



Sojojin Najeriya 'sun harbe masu zanga-zangar EndSars'
16 Nuwamba 2021

— What kind of MT content would offer best value?

Frank: Lingua Franca (*English*)

Discovery



Monitoring

&

Onward Journeys












Content creation

*An overview of **all** of WS content with a potential to travel further*

Goal:

Showcasing distinct content
across 43 World Service
teams
to facilitate and inspire
onward journeys

 <p>Give me perspective</p> <p>What does China want to achieve by teaching a 'lesson' to industrialists like Jack Ma?</p> <p>Alibaba, a key e-commerce company in industrialist Jack Ma's economic empire, was fined \$2.8 billion this April. This was the biggest fine ever imposed on a company in the economic history of the country.</p> <p>Original language: Hindi First published: Saturday 08:25</p> <p>Re-use article</p>	 <p>Diversify me</p> <p>Why are political parties opposed to the RNV? What is the politics of the BJP in the new Governor's nomination?</p> <p>Why is RNV's appointment controversial, criticizing the Liberation Party for 'doing the same job in Tamil Nadu as he did in Nagaland'?</p> <p>Original language: Tamil First published: Saturday 12:48</p> <p>Re-use article</p>	 <p>Inspire me</p> <p>What are aphorisms and how have they managed to survive in our cynical era</p> <p>In the age of social media, where brevity is valued like never before, explore the shortest of art forms.</p> <p>Original language: Spanish First published: Sunday 11:21</p> <p>Re-use article</p>
 <p>Give me perspective</p> <p>Unknown Gunmen Killed Igwe Edozieuno on Anambra State Road</p> <p>Igwe Alan Edozieuno, a former Igwe resident of Kpundo Aguleri village in Anambra State, was shot dead by gunmen.</p> <p>Original language: Igbo First published: Saturday 09:16</p> <p>Re-use article</p>	 <p>Diversify me</p> <p>Homemade recipe for chocolate eclairs</p> <p>Eclairs are a delicious and very tasty dessert. To cook them at home, you have to make an effort, but the result is worth it.</p> <p>Original language: Ukrainian First published: Sunday 07:05</p> <p>Re-use article</p>	 <p>Educate me</p> <p>Covid: What's wrong with the patient scene in the VTV movie Boundaries?</p> <p>The documentary Boundary about pregnant women infected with Covid-19 near the door of death aired on VTV is attracting public opinion but also controversial about privacy and professional ethics.</p> <p>Original language: Vietnamese First published: Saturday 09:18</p> <p>Re-use article</p>
 <p>Educate me</p> <p>The love of the cut that never stopped the cost of living</p> <p>The butchers in Addis Ababa study the past before slaughtering the bull. This background study, which aims to ensure the safety of their customers, has intensified the competition of determined traders in Addis Ababa. Unsatisfied, they fatten their own bulls and offer them to their customers. In these resolutions, from merchants to politicians, they push for the raw material.</p> <p>Original language: Amharic First published: Saturday 05:55</p> <p>Re-use article</p>	 <p>Give me perspective</p> <p>9/11: How Ramzi Al-Shaba, the facilitator of the 9/11 attacks that shook the world, was caught from Karachi?</p> <p>The arrest of Ramzi al-Shaba, the alleged facilitator of the 9/11 attacks in the United States, took place exactly one year after the attacks, on September 11, 2002, in Phase II of the Defense Housing Authority's affluent area of Karachi.</p> <p>Original language: Urdu First published: Saturday 04:45</p> <p>Re-use article</p>	 <p>Inspire me</p> <p>The historic feat of Emma Raducanu, the British teenager who won the US Open breaking several records</p> <p>British tennis player Emma Raducanu made history by winning the US Open, becoming the first tennis player to do so after entering the tournament through the qualifying round and without losing a single set.</p> <p>Original language: Spanish First published: Sunday 02:14</p> <p>Re-use article</p>

BBC Frank Dashboard

The farmers who turned a desert into a forest

Translating to Somali Change Save edit

For reference

English **Original: Spanish** Translated To: Somali

Rate this machine translation 🗣️ 😊 😐 😞 😡

Headline

Los agricultores que convirtieron un desierto en un bosque

Summary

En una región semiárida de Bahía, en el norte de Brasil, dos agricultores han conseguido crear un bosque a partir de su suelo que estaba al borde de la desertificación, gracias a un sistema agroforestal que reproduce el funcionamiento de ecosistemas originales en cada región.

Copy

Si no puedes ver el video haz clic aquí.

En una región semiárida de Bahía, dos agricultores han conseguido crear un bosque a partir de su suelo agotado que estaba al borde de la desertificación.

El suizo Ernst Gotsch y su "alumno" Nelson Araújo Filho han implementado un sistema agroforestal que reproduce el funcionamiento de los ecosistemas originales en cada región.

Translation: Somali

Copy URL Copy original CPS ID Copy entire edit

Headline

Beeraleyda oo saxaraha ka dhigay kayn

Summary

Gobolka Bahia ee waqooyiga Brazil, laba beeralay ah ayaa ku guulaystay inay kaymo ka abuuraan dhuukooda oo qarka u saarnaa inay nabaad guuraan, taas oo ay ugu wacan tahay hab-kaynta dhirta oo soo saarta shaqada hab-nololeedyada asalka ah ee gobol kasta.

Copy

Haddii aadan arki karin muuqaalka, guji halkan.

Gobolka Bahia oomane ah, laba beeraley ah ayaa ku guuleystay inay kaymo ka abuuraan carradeeda xaalufka ah oo qarka u saaran inay nabaad guurto.

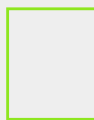
Swiss Ernst Gotsch iyo "ardaygiisa" Nelson Araújo Filho waxay hirgeliyeen nidaamka beer-forestry kaas oo soo saara shaqeynta hab-nololeedyada asalka ah ee gobol kasta.

Habkan, oo laga soo qaatay gobolka kala duwan oo Brazil iyo adduunka ah, wuxuu isku daraa wax-soo-saarka cuntada iyo dib-u-soo-celinta dhirta, soo jiidata xayawaanka duurjoogta ah waxayna ka caawisaa dib u celinta saxaraha.

Frank

into <>out of English

Filters news needs tags



Inspire me
Educate me
Divert me
Give me perspective



Update me
Keep me on trend

Next steps

Add search functions

Fine-tune filtering / affinity routes

Match with analytics



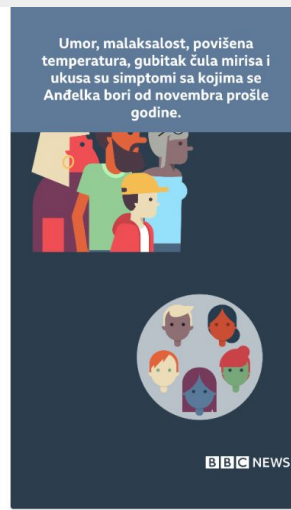
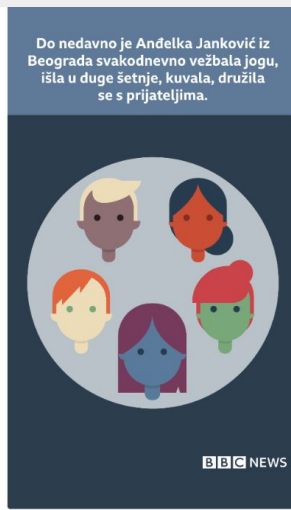
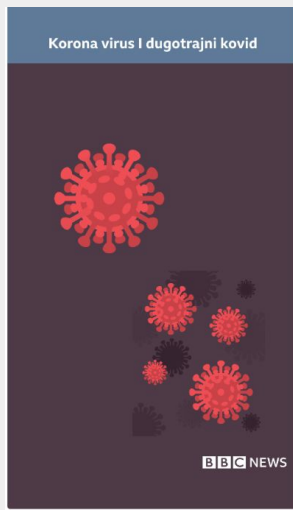
GoURMET = space to explore and experiment

How might we use machine translation as a **facilitator**
- a tool towards a solution

And not just an end in itself?



Health domain use case: Multilingual Graphical Storytelling



Graphical Storytelling (GST)

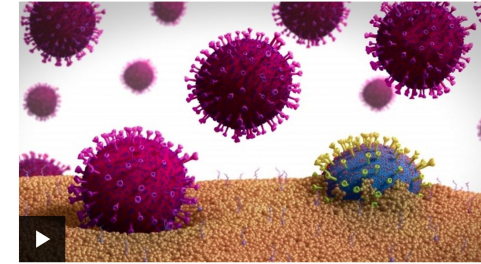
Why?

Reach underserved audiences on mobile with a more accessible story format.



Coronavirus: 'Double mutant' Covid variant found in India

© 25 March



Why do new variants of Covid-19 keep appearing? Laura Foster explains

A new "double mutant" variant of the coronavirus has been detected from samples collected in India.

Officials are checking if the variant, where two mutations come together in the same virus, may be more infectious or less affected by vaccines.

The 10,787 samples from 18 Indian states also showed up 771 cases of known variants - 736 of the UK, 34 of the South African and one Brazilian.

Officials say the variants are not linked to a spike in cases in India.

Graphical Storytelling

What?

A tool for journalists to automatically generate a graphical story from text.



Graphical Story Editor

A NEWS LABS prototype



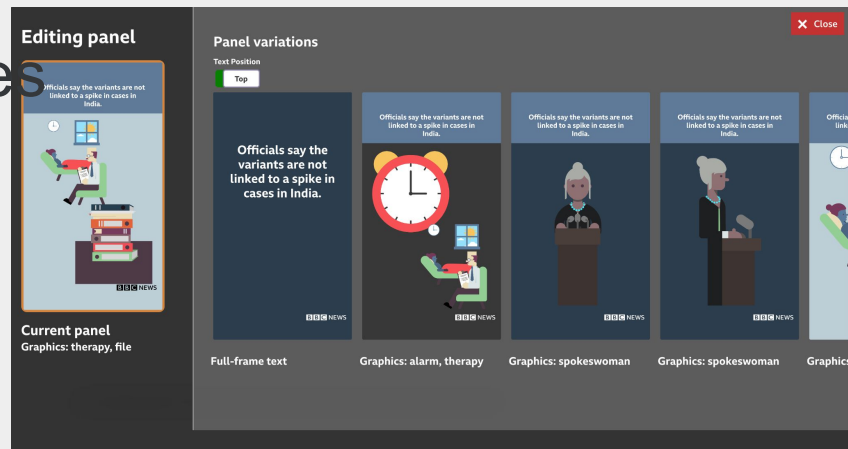
A new "double mutant" variant of the coronavirus has been detected from samples collected in India.

Officials are checking if the variant, where two mutations come together in the same virus, may be more infectious or less affected by vaccines.

Some 10,787 samples from 18 Indian states also showed up 771 cases of known variants - 736 of the UK, 34 of the South African and one Brazilian.

How?

- Entity extraction
- Matching words with images and panel types: objects, quotes, organisations, speakers
- Editing interface for panel variation



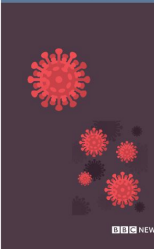
Multilingual GST

- Supports 13 languages, including Hausa and Serbian GoURMET models

BBC Graphical Story Editor A NEWS LABS prototype


DOWNLOAD STORY

Korona virus i dugotrajni kovid



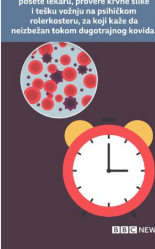
BBC NEWS

Do nedavno je Anđelka Janković iz Beograda svakodnevno vežbala jogu, ilita u duge šetnje, bunala, družila se s prijateljima.



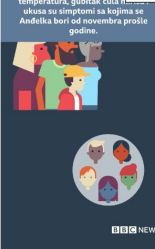
BBC NEWS

Sada se njen život sveo na redovne posete lekaru, provere krvne slike i tešku vožnju na psihickom rolerkosteru, za koji kaže da neizbezan tokom dugotrajnog kovid.




BBC NEWS

Umer, malaksalost, povisi temperatura, gubitak čula i oskud su simptomi s kojima se Anđelka bori od novembra prošle godine.



BBC NEWS

Brojni odlasci kod lekara različitih specijalista



BBC NEWS

Text in your language

Please input brief sentences in your language. Each paragraph will be displayed on a separate panel.

Korona virus: Kakvi su simptomi i šta radi telu

Virus se širio na više od 200 zemalja sveta i već mesecima traje pandemija Kovida-19.

U zemljama bivše Jugoslavije registrovano je više od 2,55 miliona slučajeva koronavirusa.

U svetu je do sada zaraženo više od 241 milion ljudi, a preminulo je više od 4,9 miliona.

Oporavilo se više od 219 miliona ljudi, pokazuju podaci Džons Hopkins univerziteta.

406/500

English translation

Your graphics are generated by translations here. Please correct if required. Corrections here do not amend text displayed in panels.

Corona virus: What are the symptoms and what does it do to the body?

The virus has spread to more than 200 countries in the world, and the Covid-19 pandemic has been under way for months.

More than 2.55 million cases of the corona virus have been registered in the countries of the former Yugoslavia.


More than 241 million people have been infected worldwide so far, and more than 4.9 million have died.

Get Panels

26 **BBC** NEWS LABS

BBC

What next?

- Building a robust tool based on learnings from Frank
 - Drafting a wider multilingual journalism proposal
 - Facilitating a comprehensive conversation on multilingual journalism
- 

Thank you!

 @BBC_News_Labs

 bbcnewslabs.co.uk

 github.com/BBC-News-Labs

 Medium.com/bbc-news-labs

