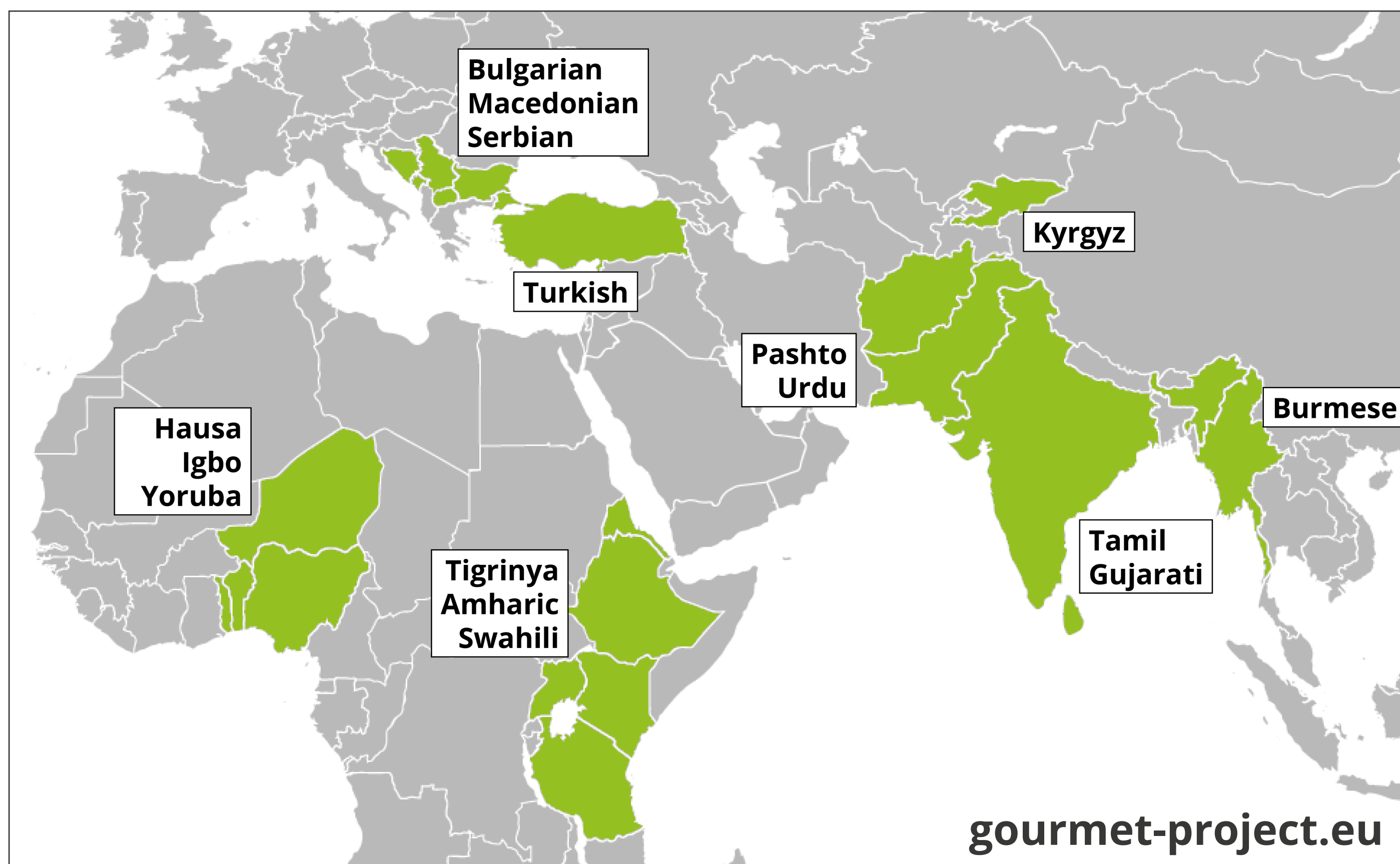


## Background



## GoURMET (Global Under-Resourced Media Translation)

is a multidisciplinary attempt to tackle challenges in low- and under-resourced machine translation for global media. It carries innovative academic research into the realm of real life deployment and rapid circulation; providing open source models as well as custom-built tools for media analysts and journalists.

## Scope

**BBC World Service** reaches 279m people every week, broadcasting in 43 languages. Most of these languages are under-resourced from a Machine Translation (MT) perspective. GoURMET's 16 languages and domains were selected iteratively over a three-year period in response to journalistic needs, strategic potential and availability of commercial alternatives.

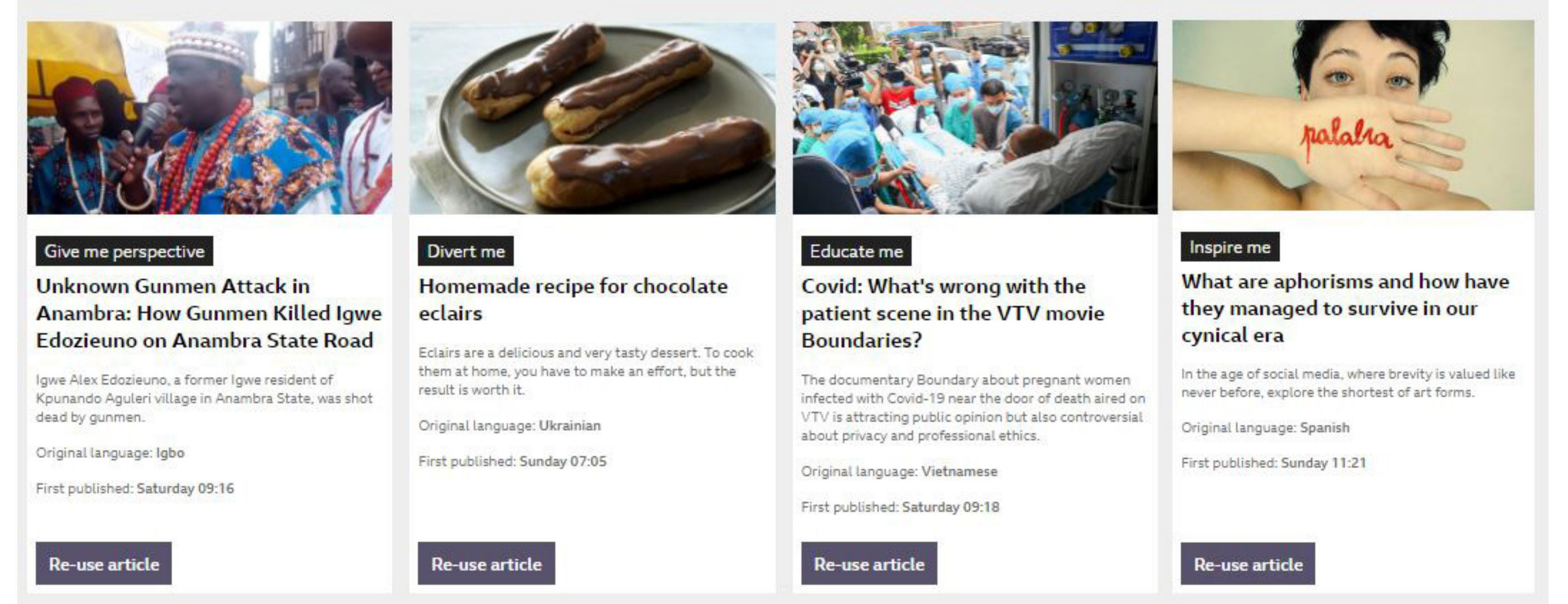
## Benefits

- Journalists are involved in every phase from language selection to evaluation, prototyping and deployment.
- The project responds to needs, learnings and emerging opportunities.
- Growing catalogue of high-quality parallel and monolingual data offers prospects for feedback loops and retraining.
- In domain-training with genuine BBC News data ensures GoURMET models follow journalistic conventions as closely as possible.



## Gains

- A swift 'surprise language' cycle focused on Pashto, a strategic language for both BBC and DW. Trained on data from both partners, the ensuing model performed better than commercial alternatives.
- Multidisciplinary events facilitated transfer of knowledge between research, development, industry & journalism spheres.
- A recent survey identified positive, exclusive stories with a longer lifecycle as an opportunity space. News Labs developed a pool and cross-recommend such content using GoURMET models to an enthusiastic reception.



- Further work includes a 'health domain' sprint when enhanced MT capabilities will be paired with automated graphics.

## Key Learnings

- Custom built in-domain models offer opportunities for under-resourced languages.
- Quality goes hand in hand with speed and scalability.
- Focus on agile prototyping allows wide-ranging test options and more insight.

## Conclusion

BBC's experience with GoURMET vindicates the value of having close multidisciplinary, iterative collaboration of researchers, developers, product owners and content creators for effective development and frictionless introduction of MT-assisted solutions in media settings.

## Challenges

- Lack or limitations of available models
- Availability of generic training data (parallel and monolingual)
- Scheduling to ensure engagement from journalism team alongside daily news cycle
- Preconceptions on usefulness of machine translated content

## Acknowledgements & Contacts

Consortium: The University of Edinburgh,  
Universitat d'Alacant, Universiteit van Amsterdam,  
Deutsche Welle, BBC (2019-2022)

info@gourmet-project.eu  
gourmet-project.eu  
@GoURMET\_MT

Funding: European Union Horizon 2020  
Research & Innovation programme  
Grant agreement No. 825299



BBC NEWS LABS

bbcnewslabs.co.uk  
@BBC\_News\_Labs

sevi.sariisik@bbc.co.uk  
@Sevi\_Sa

