

# The INTERACT Project & Crisis MT

Sharon O'Brien, Chao-Hong Liu, Andy Way (DCU)  
João Graça, André Martins, Helena Moniz (Unbabel)  
Ellie Kemp, Rebecca Petras (Translators without Borders)



[INTERACT](#)  
[International Network on](#)  
[Crisis Translation](#)



[@CrisisTrans](#)

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Marie Skłodowska-Curie grant agreement No 734211.



## Overview of Interact Project

- International Network on Crisis Translation
- EU-funded Marie Curie Networking Project (with research outputs)
- Based on the main premise that:
  - In today's age of globalisation, communication during a crisis must be multilingual and multilingual crisis communication is enabled through *translation*

# Structure

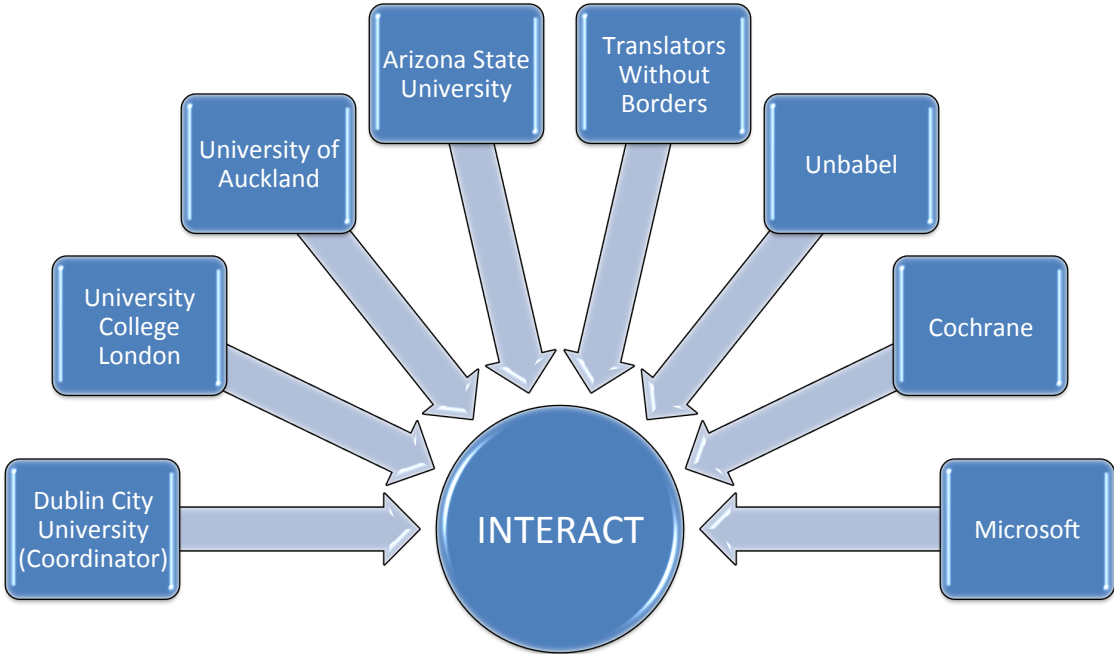
- Brief overview of project
- The role of and challenges for MT in crises
- Previous work
- What we plan to do for MT

3

## What do we mean by ‘Crisis’?

- “An event that is expected to lead to a dangerous situation, whether it is an emergency or a disaster”,  
Lighthouse Readiness Group
- Project focus:
  - Written Translation
  - Health Content

# Partners in the INTERACT Project

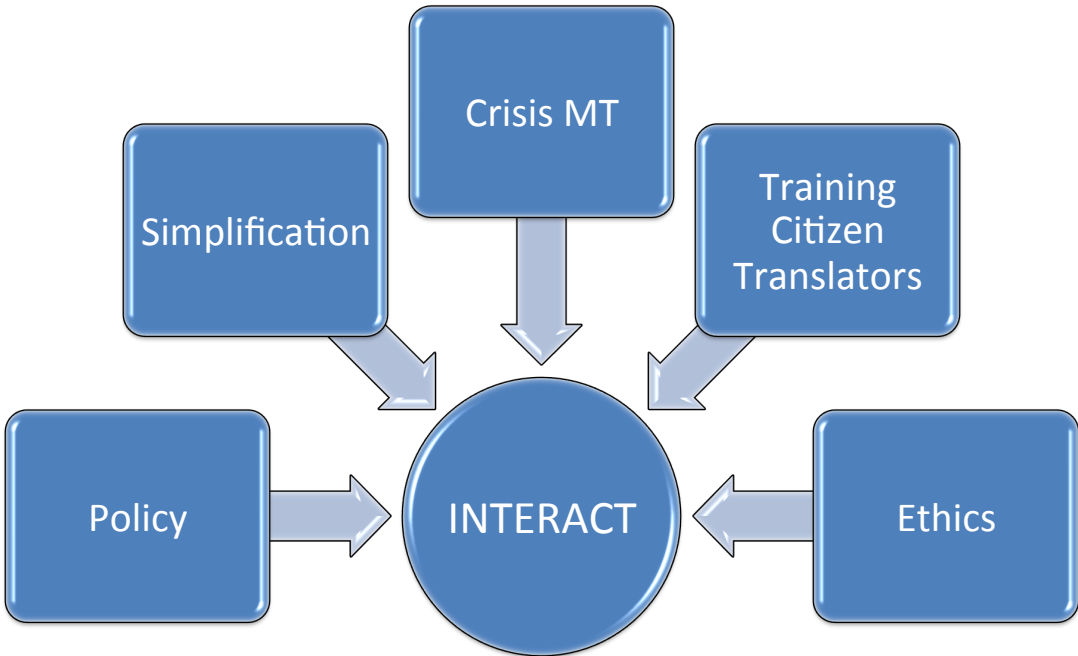


INTERACT/Crisis MT

sharon.obrien@dcu.ie

5  
@CrisisMT

# Research Work Packages

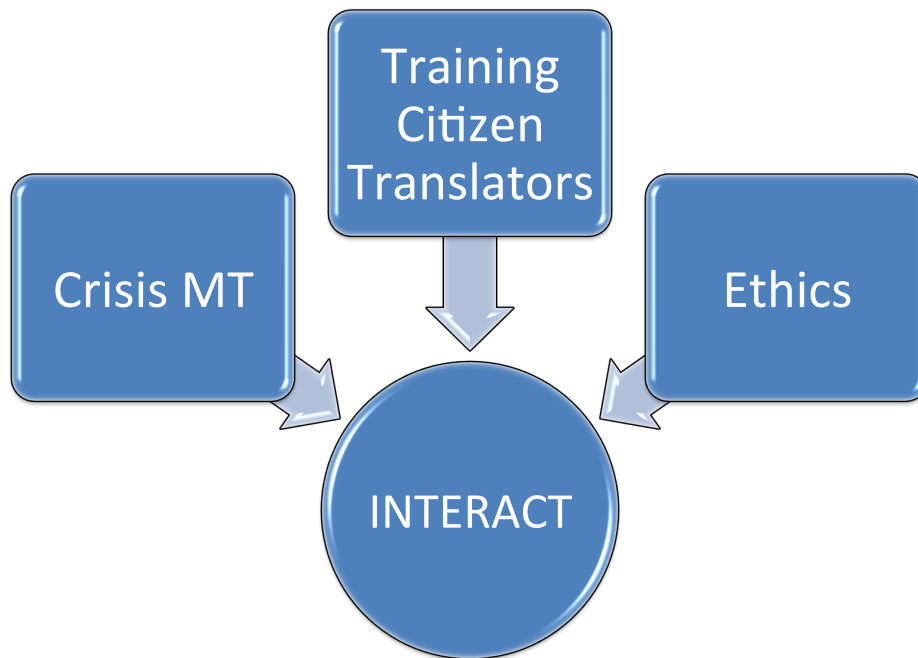


INTERACT/Crisis MT

sharon.obrien@dcu.ie

6  
@CrisisMT

# Focus For Today



INTERACT/Crisis MT

sharon.obrien@dcu.ie

7  
@CrisisMT

## Why is (Machine) Translation Important in a Crisis?

- Clear, accurate, timely information is essential in a crisis
  - e.g. Seeger 2006; Fischer 2008; World Health Organisation 2012; Infoasaid 2012; Santos Hernández and Morrow 2013
- Greater cooperation between humanitarian agencies and linguistic volunteers is required (Harvard Humanitarian Initiative 2011)
- but...
- Little to no recognition of the fact that those in need of information may not speak the dominant 'response' language

INTERACT/Crisis MT

sharon.obrien@dcu.ie

@CrisisMT 8

# Nature of Crises

- May have sudden onset
- Unpredictable language combinations and information needs
- Unpredictable duration
- Highly stressful
- Lives are at risk
  
- But also consider: the 4Rs of Disaster Management:
  - Risk
  - Response
  - Recovery
  - Resilience

i.e. Translation is not just required during the *response stage*

## Examples of Previous & Ongoing Work

- MT in response to the Haiti Earthquake (Lewis 2010)
  - “Cookbook” for SMT in crisis situations (Lewis et al. 2011)
- DARPA’s Lorelei (Low Resource Languages for Emergent Incidents) project
- TWB’s Rule-Based MT systems
  - English-Kurmanji, English-Sorani via Apertium

# Challenges for Crisis MT

- Unknown time of occurrence (no training in advance?)
- Unknown language pairs
- (Often) low resource languages
- (Often) no/little parallel data for affected languages
- (Sometimes) highly specialised (communicating risk of disease, nuclear threat etc.)
- (Sometimes) no/low power and Internet connections
- Text to Speech may be required too

## Pivoting as Potential Solution?

- E.g. Arabic-speaking refugees arrive on a Greek Island. Responders speak Greek and limited English. Refugees speak Arabic, some have limited English, and no Greek
- For emergency response purposes, e.g. ascertaining the state of health of refugees,
  - we need Arabic < > Greek translation urgently
- We do not have sufficient translators/ interpreters
- We have Arabic-to-English and English-to-Greek Engines, but no Arabic-to-Greek engines

# Pivoting as Potential Solution

- Our questions: Could we use:
  - Arabic < > English < > Greek in this situation?
  - What is the quality like?
  - What impacts the quality?

## Approaches to Pivoting

- Naïve Approach (Utiyama & Isahara 2007)
  - Translate from A to B
  - Use B as source to translate to C

# Approaches to Pivoting

- Interpolated Direct Engine (Wu and Wang 2007)
  - Applies only to PBMT
  - A-to-C phrase table is derived from available trained A-to-B and B-to-C models
  - Combined with language model available for C to build direct engine from A-to-C

INTERACT/Crisis MT

sharon.obrien@dcu.ie

@CrisisMT 15

# Approaches to Pivoting

- Neural Interlingua Approach (Johnson et al. 2016)
  - ‘One’ Neural Network is trained with A-to-B and B-to-C sentence pairs
  - NN is used to translate A-to-C even though no A-to-C sentence pairs are used in the training
  - Performance improved if small amount of A-to-C sentences are used in training

INTERACT/Crisis MT

sharon.obrien@dcu.ie

@CrisisMT 16



# Our Pivot Triplets

1. Greek < > English < > Arabic
2. German < > English < > Arabic
3. French < > English < > Swahili

The task will involve identifying relevant training resources, building sample crisis MT engines for health content, and evaluating the results using standard MT evaluation techniques (automatic and human evaluation metrics).

## The Human Factor

- (Citizen) Translators
- End Users, e.g. First Responders
- How do we train these people to work with MT in crises?

# The Human Factor

- Development of training materials for Citizen Translators who volunteer in Crisis Scenarios
- One specific focus will be post-editing
  - E.g. Through the Unbabel crowdsourcing platform and TWB's and Cochrane's Networks
- Train, evaluate, re-design, train, evaluate...
- Peer review via the crowd (e.g. using Unbabel's online MQM annotation tool)
- Train the trainer
- Training in support resources, e.g. Slándáil, ReliefWeb, TWB harvest of terminology from Sphere Handbook, etc.

# The Human Factor

- Focus on end users too:
  - Ethical and informed use of MT for crisis scenarios
  - Use of Quality Estimation for *triaging* MT output so unreliable translation is not even seen by post-editors/end users

# To Conclude

- Follow our progress on Twitter:  
[@CrisisTranslation](https://twitter.com/CrisisTranslation)