



# INTERACT Deliverable 8.3: Panel on Crisis MT

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.



# Ethical considerations on the use of machine translation and crowdsourcing in crises

Carla Parra Escartín  
Unbabel

Helena Moniz  
INESC-ID/Unbabel

# Translation as a Human Right

According to the Universal Declaration of Human Rights:

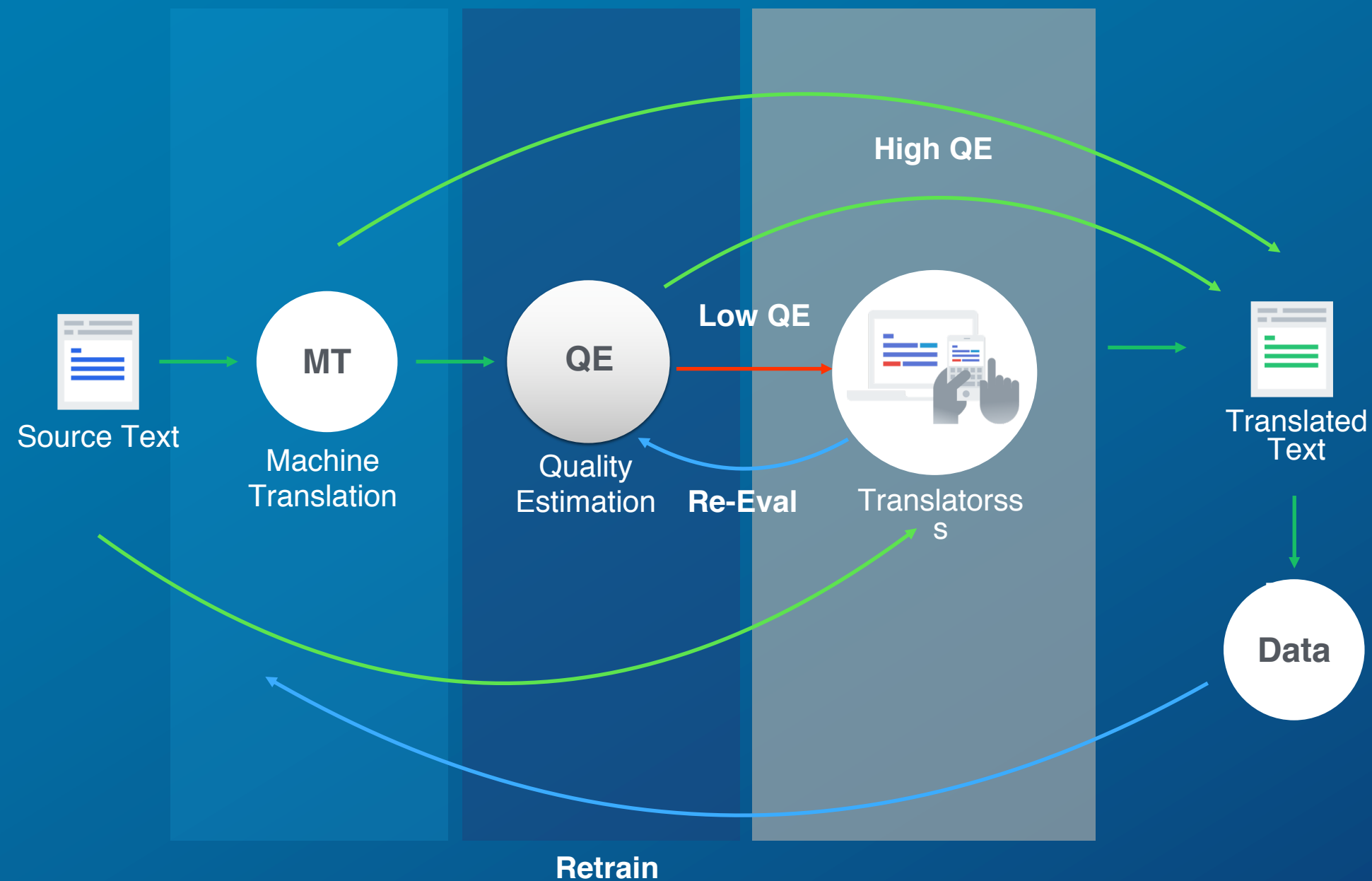
**“Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status. Furthermore, no distinction shall be made on the basis of the political, jurisdictional or international status of the country or territory to which a person belongs, whether it be independent, trust, non-self-governing or under any other limitation of sovereignty.”**  
(Article 2)

# Let us imagine...

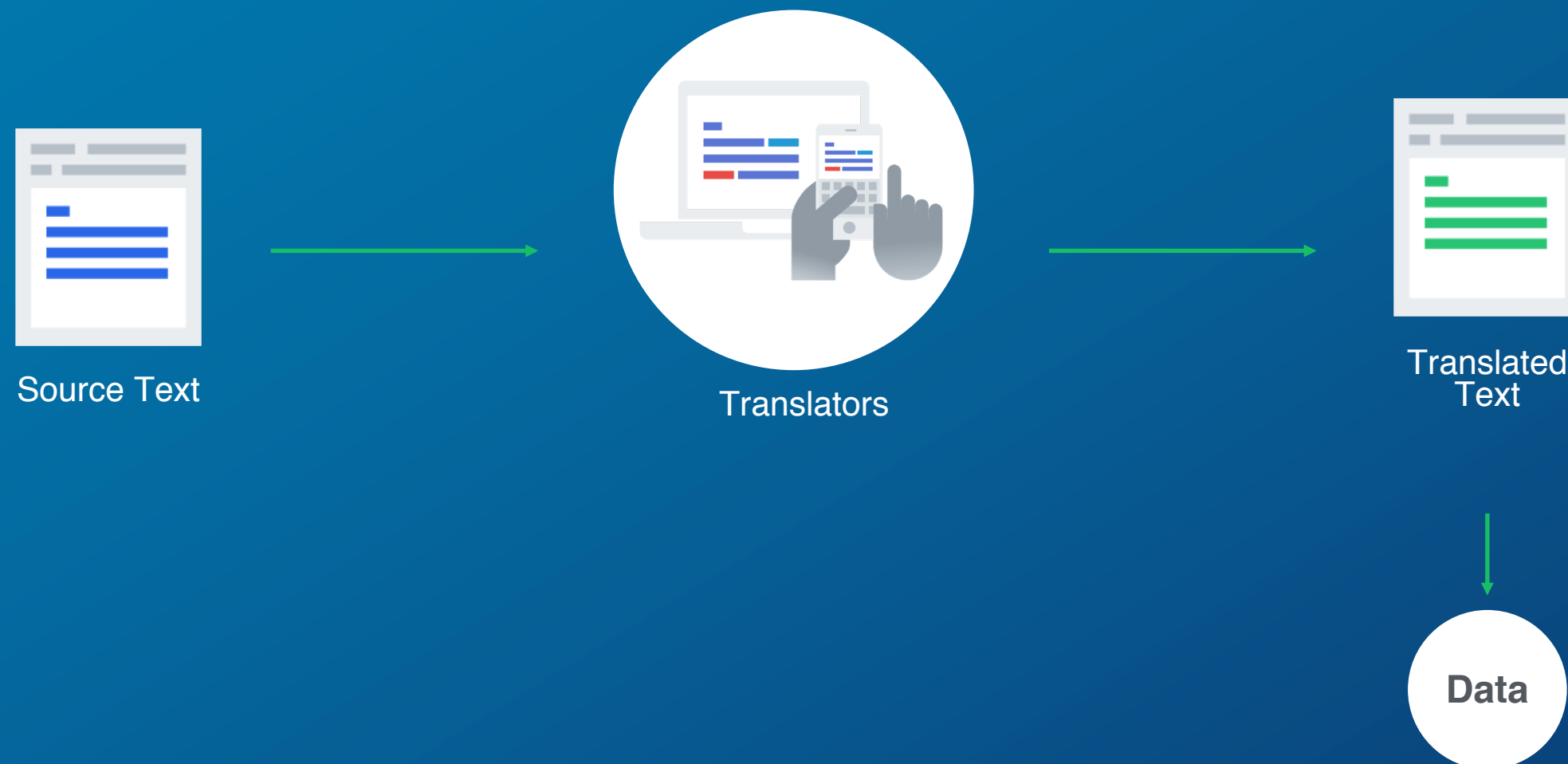
- We are in a refugee camp and a doctor needs to give assistance to refugees but they don't share a common language
  - We are visiting a country that suddenly gets hit by a tsunami
  - We suddenly experience major new volcanic activity in the Auckland (NZ) region
  - We are tourists in a city where there is a terrorist attack
- Who are the actors?
  - Who needs translations?
  - Who translates?
  - How do we ensure Quality?
  - What ethic concerns?

workflows involving machine translation are

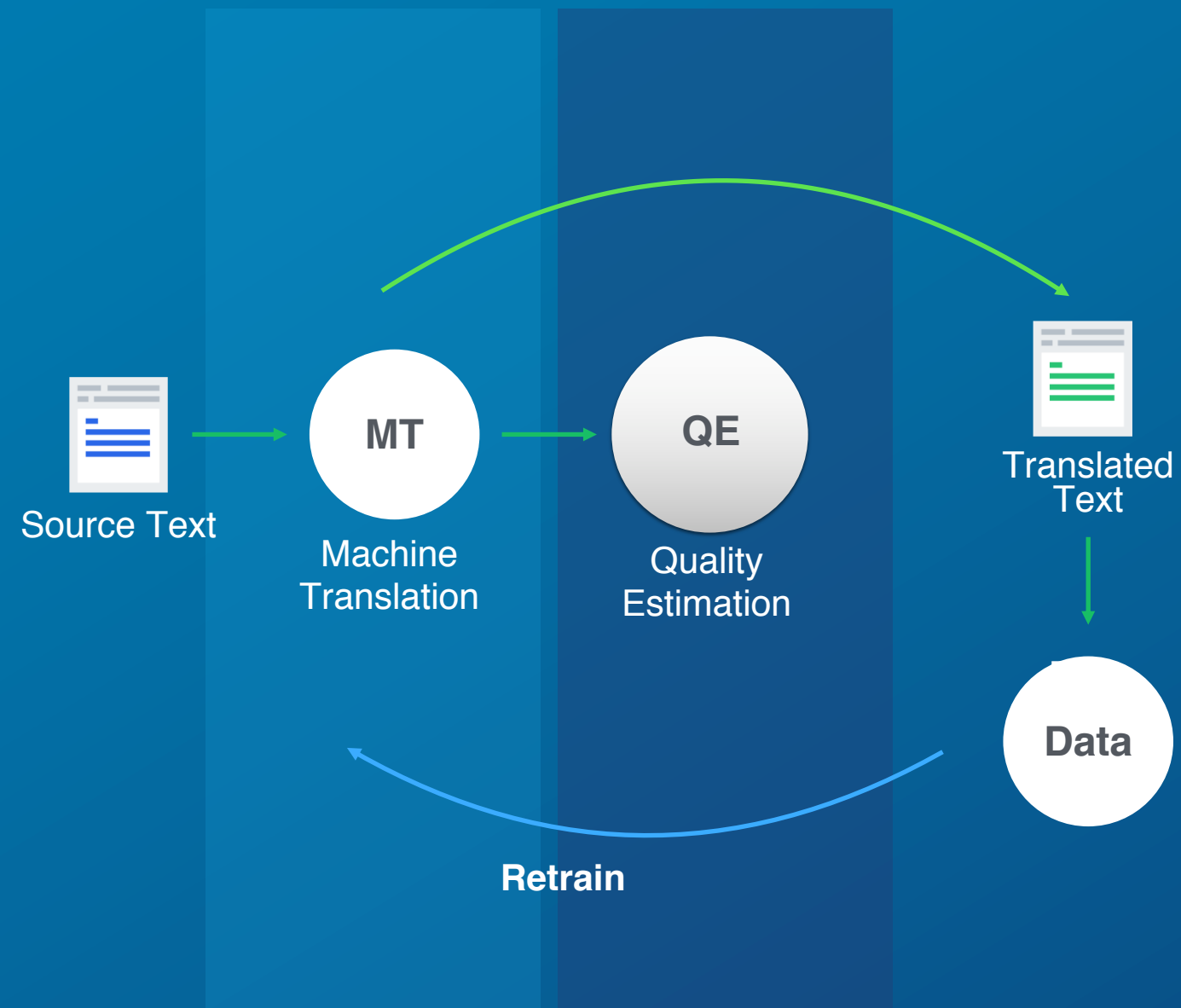
# Holistic view of current workflows



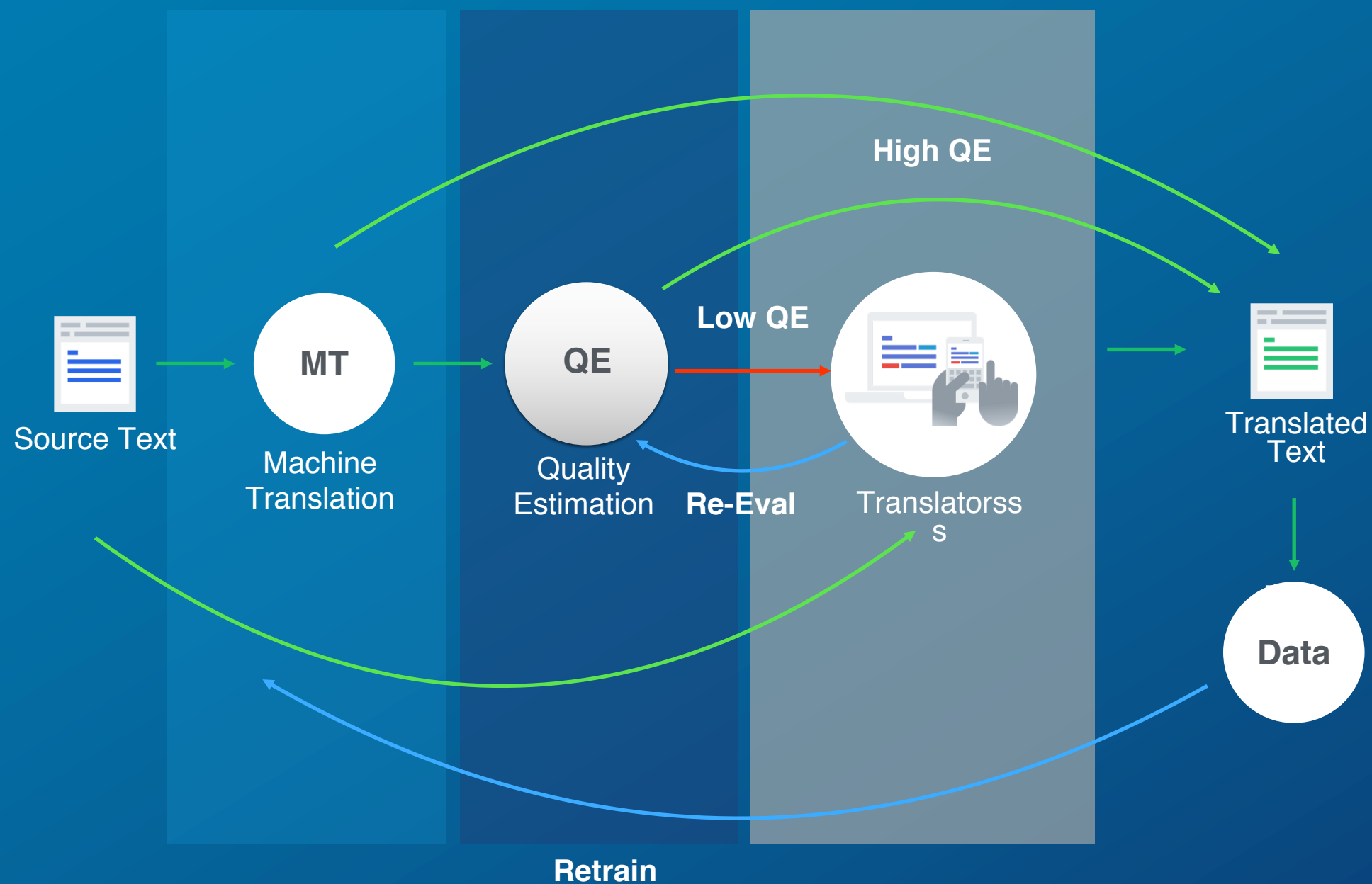
# Flow 1: (Technology Assisted) Translation with



# Workflow 2: Machine translation



# Workflow 3: Machine translation and post-editing



# Ethical considerations on Data, AI, and human actors in crisis situations

# Core questions

- How can we ensure that citizens are safe?
- How do we ensure privacy?
- Which data are we going to use?
- How do we deal with sensitive data?
- Can we keep data from this crisis to use it in the next?
- What about quality?
- How can technologies assist human beings taking all the above into consideration?

# Ethical implications on data and crisis

- Artificial Intelligence seen as the Fourth Industrial Revolution supported on the pillar of the availability of Big Data
- Citizens' data available on the web or dedicated servers owned by public and/or private institutions – we term this Big (our) Data

# implications on data and crisis: several init

- **The IEEE Global Initiative**
- **EU-funded projects** (e.g., SATORI and SIENNA) to discuss and frame ethical considerations
- General Data Protection Regulation (**GDPR** - The European Parliament and The European Council 2016)

# Ethical implications and data

- Privacy issues — Sensitive nature of data in a crisis
- Robust anonymisation
- Data shared in different devices
- Distinct stakeholders
- Distinct legislation across countries
- Quality of the data
- Ownership of the data and curation
- Reusability of the data

# AI principles and implications for crises

# AI principles

- The IEEE Global Initiative — Ethically aligned designed
- The four principles:
  - (1) Human Benefit,
  - (2) Responsibility,
  - (3) Transparency, and
  - (4) Education and awareness

# AI principles

- The principles are based on three premises:
  - (1) they should embody the highest ideals of human rights,
  - (2) they should prioritise the maximum benefit to humanity and the natural environment, and
  - (3) they should mitigate risks and negative impacts as AI evolve as socio-technical systems.

# (1) Human benefit

## **Development of AI systems that do not infringe human rights**

- MT technology and crowdsourcing techniques used for disaster relief should **observe the Universal Declaration of Human Rights** (UN General Assembly 1948)
- The **infrastructure** deployed must be **secure**, and all **steps** should be **traceable**.

## (2) Responsibility

### Assuring that AI systems are accountable

- Legislation that covers AI for disaster situations by:
  1. Taking into account the **different cultures and stakeholders** involved in a crisis situation
  2. Integrating all stakeholders into a **common translation framework** and ensuring they all work together and are **aware** of the translation workflow
  3. Duly **documenting** the setting up the **infrastructure** for the translation workflow(s).

# (3) Transparency

**AI systems that are clear and explainable, to build public confidence in the technology**

- **MT output or QE system clear and understandable** to all the users
- **Relevant stakeholders** should be **involved** in the process of **establishing the thresholds** and scores to be used (for QE, for instance) and assess them in terms of quality risks

# (4) Education and awareness

**Extending the benefits of AI while minimising the risks of technologies being misused**

- **All stakeholders** involved in the translation workflow need to be trained:
  - How MT systems work
  - How the infrastructure has been implemented to allow for crowdsourcing translation

# Conclusions

- The infrastructure deployed must be secure
- All steps should be traceable:
  - Who was responsible for a translation error?
  - Where in the MT system a translation issue that causes problems (e.g. mistranslations, omissions, additions, etc.) is generated?

# Conclusions

- The best of AI systems augmented by human editing will allow to **assist human beings** in diverse crisis situations.
- To ensure the balance between AI systems and human beings the **IEEE principles** should be applied.
- Our overall recommendation would be to **engage professional translators** whenever possible, and use their professional expertise as coordinators of the translation task and quality assurers.



# Thank you!

[helena@unbabel.com](mailto:helena@unbabel.com) & [carla.parra@unbabel.com](mailto:carla.parra@unbabel.com)

from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska